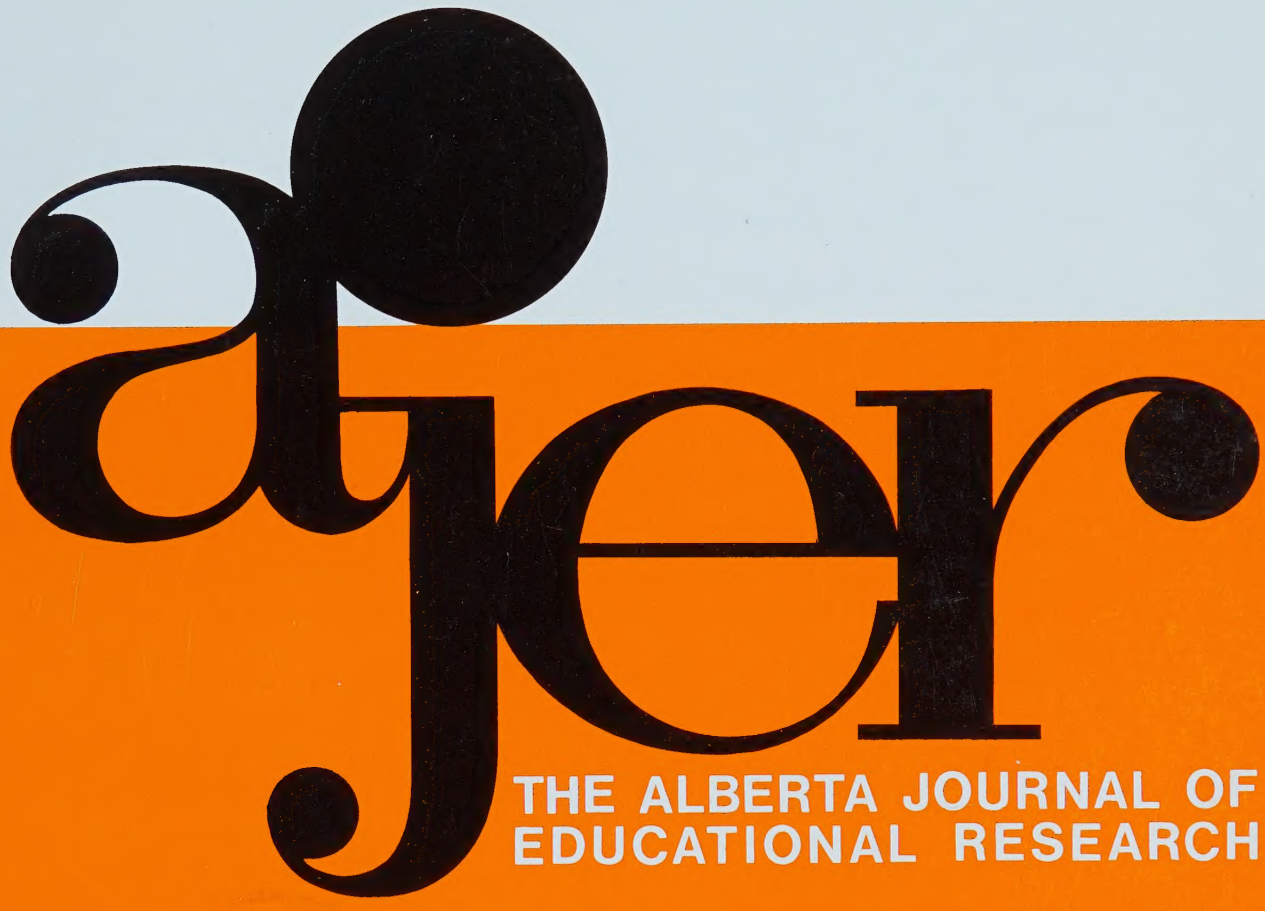


THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXVII NUMBER 1 MARCH 1991

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Thomas O. Maguire

Consulting Editors

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

Janet Donald
McGill University

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Lilian Katz
University of Illinois

Les McLean
Ontario Institute for Studies in Education

John McLeod
University of Saskatchewan

Linda Phillips
Memorial University of Newfoundland

Lauran Sandals
University of Calgary

Keith Sullivan
Dalhousie University

Phil Winne
Simon Fraser University

Ruth Wright
University of Ottawa

Faculty Advisory Committee

Harvey Zingle (Chair)
Fern Snart
Erwin Miklos
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Second Class Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXVII, Number 1, March 1991

EDITORIAL	1	A Note of Thanks
JAMES COVERT LEONARD WILLIAMS WILLIAM KENNEDY	3	Some Perceived Professional Needs of Beginning Teachers in Newfoundland
MARGARET R. OLSON	19	Supportive Growth Experiences of Beginning Teachers
CLAIRE F. STAAB	31	Teachers' Practices With Regard to Oral Language
FRANK H. ECHOLS, Jr. CHARLES S. UNGERLEIDER CAROL LA BAR	49	Students' Knowledge of the Charter of Rights and Freedoms and Their Agreement With Its Provisions
NELLY McEWEN HAU CHOW	65	Issues in Implementing Indicator Systems
KATHLEEN OBERLE	87	Paradigm Wars: Who's Fighting; Who's Winning?

ISSN 0002-4805

Acknowledgment

The review process is central to the quality of this journal. We thank the following people who reviewed manuscripts in 1990. Their thoughtful critiques contribute to the advancement of theory, research, and practice in education.

Kelsey, G.
Ahola-Sidaway, J.
Bacchus, M.K.
Bachor, D.G.
Ball, C.
Beauchamp, L.S.
Bérard, R.
Birnie, H.
Blakey, J.
Boersma, F.J.
Bogdan, D.
Browne, M.P.
Bryce, R.
Bullcock, J.
Burnett, D.
Butt, R.
Callan, E.
Campbell, D.
Carson, T.R.
Chamberlin, C.
Clifton, R.
Crocker, R.
Edge, D.
Edwards, J.
Elliot, M.
Enns, R.
Erickson, G.
Everett-Turner, L.
Goldsmith, C.
Greene, M.
Greenfield, T.
Hare, W.
Hart, P.
Hattie, J.
Hoge, R.
Holdaway, E.
Holmes, M.
Housego, B.E.G.
Hunka, S.M.
Irwin, R.
Jackson, M.
Jefferson, A.
Janzen, H.L.
Juliebo, M.F.
Korteweg, L.
Krentz, C.
LaRocque, L.
Leblanc, D.
Leong, C.
Leroux, J.
MacIvor, D.
MacRury, K.
Madak, P.
Malicky, G.
Mansfield, B.
Marcotte, P.
McIntyre, I.
McLean, L.
McNay, M.
McNeill, J.
Miklos, E.
Muir, W.
Murphy, P.
Nagy, P.
Newton, E.
Norman, C.A.
Norris, S.
Ollila, L.
Olson, A.T.
O'Reilly, R.
Pearson, A.
Phillips, L.
Prock, L.
Pyryt, M.
Ratzlaff, H.
Ragsdale, R.
Richards, M.
Richer, E.
Rogers, T.
Rowell, P.
Russell, T.
Ryan, A.
Sanche, R.
Sande, D.
Shapiro, B.
Shore, B.
Sigurdson, S.E.
Snart, F.D.
Staab, C.
Sullivan, K.
Thomas, A.
Traub, R.E.
Uhl, N.
Ungerleider, C.
Vaala, L.
Van Hesteren, D.
Wahlstrom, M.
Whitehead, L.
Wiener, J.
Wilde, W.
Wilgosh, L.
Wilson, K.
Wood, A.B.
Wood, J.D.
Wright, R.
Young, B.
Young, Jean
Young, John

The Alberta Journal of Educational Research *Vol. XXXVII, No. 1, March 1991, 1-1*

EDITORIAL

A Note of Thanks

Each year at the beginning of the March issue, the assistance of the reviewers is acknowledged. Their work constitutes a significant part of the research process, but the listing of their names hardly does justice to the effort that they put into their task. I try to draw expertise through two sources: the *CSSE Directory* and the contacts of the members of the editorial board. From the quality of the reviews that I receive, I believe that in most cases I have been successful in locating people who possess both insight and sensitivity. They labor for the benefit of the academy, and they do not accumulate a lot of points on the salary and promotions ladder for it. Nevertheless, I would like to remind them all that it is an essential part of the research process, and to thank them for making my editorial task possible.

It is often the case that I receive manuscripts from new researchers. Sometimes these are not in perfect form. Yet time after time I find that the referees recognize their responsibility to new scholars and take great pains to provide constructive suggestions for improvement. I seldom encounter destructive comments, and I take this to be a positive comment on the motives of the membership of the Canadian educational research community.

The editorial board has been particularly helpful in providing commentary on difficult manuscripts. Although there may be some "glory" in having your name listed on the inside front cover four times a year, it hardly compensates for the amount of time that they provide. Moreover, being on the cover does not mean that they are not called upon for regular reviews. Ours is a small community and we need all the help that we can get.

Finally, as we begin this 37th year of *AJER*, I would like to thank the contributors—both those whose work appears and those who are not as successful. The quality of most of the submissions is very good, and there simply is not enough room for all of them. I have tried to keep a balance among research areas, and as a consequence there are some manuscripts that are treated more critically than others simply because there is a large pool to draw from in that particular field and I do not want to make the journal become a journal of teaching, or a journal of measurement, or a journal of administration. Overall, however, I hope that I have erred on the side of being too inclusive. Publication is an important part of the research process and some of the newer fields need the sort of nurturing that *AJER* can provide in order to get their case in front of the research community.

T.O.M.


Digitized by the Internet Archive
in 2024 with funding from
University of Alberta Library

https://archive.org/details/ajer_1991

The Alberta Journal of Educational Research *Vol. XXVII, No. 1, March 1991, 3-17*

JAMES COVERT

LEONARD WILLIAMS

and

WILLIAM KENNEDY

Memorial University, Newfoundland

Some Perceived Professional Needs of Beginning Teachers in Newfoundland

As a first step in the induction process, a survey of professional needs was sent to all beginning teachers in the province of Newfoundland. The respondents had several concerns that were similar to those reported in the induction literature, but because nearly 75% begin their careers in rural areas, they may have identified some unique concerns for beginning teachers in most rural communities. Beginning teachers were concerned with the difficulties of dealing with a multigrade classroom and its wide variety of courses. The tight job market in Newfoundland and the resultant probability of becoming employed only in rural schools led to concerns about entry through the substitute route and the legal rights of teachers being dismissed while on probation. There was a wider interest in the legal rights of teachers and students expressed by the respondents. The increasingly litigious nature of the society and the high profile of child physical and sexual abuse may weigh heavily on first-year teachers. This study indicates that induction in rural areas may be much more difficult and complex than in more populated areas. Some innovative programs are suggested to solve the unique problems of induction in rural schools.

Introduction

There is such a growing awareness of the need for programs for beginning teachers in Canada that the 1990s may be known as the decade of induction. Even though there are no official province-wide induction programs in place in Canada there is a great interest in the concept, and several attempts have been made at different educational levels to institute programs that are of assistance to beginning teachers. Ontario may have the most ambitious set of programs, with school boards and teacher education institutions cooperating in various ways to ease the transition from student to teacher. McNay and Cole (1989) and Cole and McNay (1989) have written about induction devel-

James Covert is a professor of educational foundations with a research interest in the professionalization of teachers.

Leonard Williams is a professor of education and the former Director of the Division of Student Teaching. He is currently chairing a Royal Commission on Education for the Newfoundland government.

William Kennedy came to Memorial University from the Newfoundland Department of Education as an associate professor with the Student Teaching Division and was recently appointed Director of Student Placement.

opments in Ontario and are the editors of *Inducere*, a newsletter published by the University of Western Ontario dealing exclusively with induction.

Much of the impetus for the induction programs in Ontario has come from the influential Fullan and Connelly Report (1987) which recommends

> That a two-year period of induction become mandatory for all teacher initiates who have completed a one-year pre-service teacher education program or a concurrent pre-service program. We propose that the teaching load during the first year of induction be no more than 60 percent and no more than 80 percent in the second year." (p. 76)

The University of British Columbia has taken the initiative by establishing several partnerships with local school boards through the Centre for the Study of Teacher Education. The Centre examines different approaches to induction through a vigorous research program and publishes a newsletter dealing with various educational issues including induction. It is interesting to note that these two provinces have the most to gain from retaining beginning teachers because of the teacher shortages that they are facing in large urban centers in both provinces.

Two recent studies in the United States indicate that retention of beginning teachers may be a problem. Huling-Austin (1986) reports that as many as 15% of beginning teachers leave teaching during their first year, and as many as 50% may leave during their first five years of teaching. Perhaps of greater concern is the finding by Varah, Theune, and Parker (1986) that beginning teachers with the highest university academic standing are the most likely to leave. This is of particular interest because academic performance is an important criterion used for admission into most teacher education faculties in Canada.

The reasons for undertaking this study came from a different set of concerns. From research in previous years it has become apparent that the vast majority of the first-year teachers employed in Newfoundland have begun their careers in rural areas. The persistent decline in student enrollment through the '70s and '80s has led to a difficult job market for teachers. In addition to this situation, in 1987 the provincial government sought a review of the teacher education program in the province. The subsequent report, *Focusing our Future* (1988), made two recommendations concerning the induction of teachers.

> Recommendation 44: That school boards provide an induction program for beginning teachers, extending over the two-year probationary period.
>
> Recommendation 45: That a provincial induction program plan be developed jointly by the department of Education , the Newfoundland Teachers Association, The Faculty of Education and the Newfoundland and Labrador School Trustees' Association. (p. 54)

Considering the national interest in induction and the local encouragement for action, this study seem to be a logical first step. Although there is much interest in induction in Canada, most of the Canadian researchers indicate a dearth of Canadian research on the subject. Grimmett (1984) suggests that empirical information for making informed decisions about teaching at the preservice and induction levels is virtually nonexistent.

Wideen and Holborn (1986), when talking about change in faculties of education in Canada, failed to even mention induction, though one of the major movements discussed was lengthening practical experience and integrating theory and practice. There is little specific information available about the needs of beginning teachers in Canada, but there is little to indicate that these needs would vary much from those that have been well documented in the United States (Snow, 1988).

Griffin (1985) is encouraged by the research activity that has taken place in the past five years in the United States but notes that this is a relatively recent occurrence.

> McDonald (1980) noted: We have little information about how teachers pass through the transition stage, other than to know that some apparently do so successfully, some do not. We have no detailed information on how those people who master the transition period do so. Nor do we have information on how different kinds of assistance or help directly or indirectly influence the teachers' successful mastery of the induction period. (p. 44)

Huling-Austin (1986) also indicates the rapid growth of induction programs in the last decade.

> A survey of the educational literature in the mid-1970s would have yielded few articles about the problems of teacher induction and programs to assist beginning teachers. Of those few who were writing about teacher induction, most were from other countries, namely Great Britain and Australia.... A survey of the current literature indicates that teacher induction is a topic that is growing in popularity. In fact, several major journals have devoted entire issues to the topic including *Educational Leadership* (November, 1985), *Journal of Teacher Education* (January-February, 1986), *Kappa Delta Pi Record* (July-August, 1986) and *Action in Teacher Education* (Winter, 1987). (pp. 3-4)

These studies indicate that induction is a complex activity with many different possible programs that might be applicable. The most prevalent kind of program is some form of mentoring, where a more senior teacher provides support for the beginner in a variety of ways. Obviously these programs are only as good as the mentor teachers, so much of the research has been directed toward the selection and development of mentor teachers. Perhaps the recent article by Zimpher and Rieger (1988) is indicative of the high quality of work being done on this aspect of induction. The influence of the school climate on teacher induction is well explained by Bullough, Knowles, and Crow (1989).

Perhaps the newest trend in teacher induction is the development of the clinical teaching schools. These are a cross between the teaching hospitals that train doctors and nurses and the Dewey-inspired laboratory schools that have all but faded from the educational landscape. Darling-Hammond (1989) summarizes them this way:

> Such schools, in their infancy in more than a dozen cities around the country, are intended—like teaching hospitals—to model state-of-the-art practice while simultaneously refining and spreading it. They are to be places where experts train and socialize novices, where research and theories are trans-

> lated into practice, and where practical knowledge is translated in turn into research and theory. (p. 59)

Although all these are important considerations, the programs that are being initiated in Canada are much less well developed and are in need of some basic needs assessment to determine if the problems facing schools in Canada are unique. This study indicates that perhaps there are certain differences that ought to be considered, especially when dealing with beginning teachers in rural areas that do not have access to all the amenities available to teachers in urban areas.

Questionnaire Development

In the American research concerning the needs of beginning teachers, several topics appear frequently. McEvoy and Morehead (1987) indicate that the needs listed most frequently are classroom management, instructional strategies, student evaluation, and the location and development of learning materials. Hawk (1987) adds that interacting with parents and offering emotional support for new teachers are two items that are perceived to be of importance but to a lesser degree. Odell (1986) summarizes induction needs this way:

> The two primary needs are (a) to obtain fundamental information about the school district and (b) to obtain resources and materials pertinent to the information to be taught. As experience in the school increases, and the need for system information decreases, the new teachers appear to demand more help around teaching strategies and the instructional process.... The present data show that emotional support is important, but like classroom management support, it is less needed than is help in obtaining resources and materials or in applying a given teaching strategy. (p. 29)

The American studies dealing with the needs of beginning teachers have been quite consistent during the period of induction investigation. Howey (1988) indicates some of these concerns for American teachers.

> Teacher preparation institutions nationally (Loadman, 1983, cited in Zimpher, 1986) have revealed some common concerns of these beginning teachers, among them: discipline, communication skills, lesson and curricular planning, student evaluation, as well as other skills beginning teachers claim they never acquired or that were inadequately treated in their programs (deVoss, 1978, 1979, 1980; Katz, Raths, Mohanty, Kurachi, & Irving, 1981, cited in Zimpher 1986, p. 17). (p. 209)

The eight topics selected for this study were those consistently and most frequently identified in the American research literature that were judged to be most pertinent to the Newfoundland situation. A section on the questionnaire was provided for respondents to comment and list any other concerns encountered during their first year of teaching. Respondents were asked to assign a relative rating to each topic in accordance with their perception of its importance. They were offered a three-point scale: the most important topics were to be rated as one, the next important as two, and the least important as three. Kozioh and Burns (1986) indicate several concerns about teacher self reporting, but this seems to be the most appropriate means of assessing the needs of beginning teachers for the purpose of induction. Schlechty

(1985) argues that teacher induction is more like an inservice program than preservice teacher education, so it is important to consider their concerns when designing the support program for beginning teachers. It would be extremely difficult to impose an induction program on teachers if it failed to meet their needs, especially if there is no compulsion to attend.

Because of the relatively small size of the beginning teacher force in Newfoundland, it was possible to survey the entire complement of first-year inexperienced teachers. The questionnaire was administered to 118 beginning primary, elementary, and secondary teachers who were at the first incremental point of the Newfoundland provincial teachers' salary scale as of September 1988. Because beginning teachers are on a two-year probationary period and might be concerned that this questionnaire could be used for evaluative purposes, little background information was sought. The teachers were not differentiated on the basis of their teaching level or their sex. The questionnaire was sent to the school address of each teacher, so it was possible to determine whether their schools were in rural or urban areas.

The questionnaire was sent to the beginning teachers in March 1989. Teachers were then in their sixth month of teaching. It was thought that after six months, beginning teachers would have had sufficient time to become accustomed to their teaching situation and to be able to appreciate the topics under consideration. Approximately 50% (57 out of 118) of the respondents returned the questionnaire. Because of the time of the year and the busy schedule of the first-year teacher, this was considered to be an acceptable return.

In part the survey instructions read, "In order to establish the most helpful in-service sessions for next year's beginning teachers, would you please rate the following activities." Therefore, the topics under consideration were presented as suggestions for inservice activities. Beginning teachers were asked to make a judgment about the importance of each of these topics in accordance with their perception of the importance of this concern during their first year of teaching. The topics were further explained according to the following instructions. "We have listed two examples that might be used to help identify each of the activities."

Beginning teachers could rate each of the items as one (most important), two (less important), or three (least important) to each of the topics. They were permitted to rate more than one response as most important; in fact they could rate all of the questions as of greatest importance. If this were the case, any one topic could receive a score of 57, a rating of most important from each of the responding teachers. In this way the topics were rated in relation to their perceived importance to each teacher. In an effort to determine the overall importance of each of the topics, the individual responses of the teachers were simply recorded in their respective category of importance. The individual responses of the 57 teachers to each of the eight items are included in Table 1. Not all of the teachers responded to each of the items, so three of the eight rows fail to total 57.

Table 1
Rating of all Responses to Questionnaire Items
N = 57

Topic	Most Important	Important	Least Important
Classroom Discipline	49	7	1
Student Interaction	36	18	1
Instructional Management	34	19	4
Resource Utilization	31	20	4
Personal Motivation	28	22	0
Student Evaluation	28	20	9
Public Relations	22	26	9
Communication Skills	21	15	21

Findings

Table 1 indicates how the beginning teachers ranked the eight topics. Classroom discipline is clearly the topic of greatest importance and communication skills is the topic of least importance. Following discipline, the next three topics of interest to the teachers concern the mechanics of classroom operation. These three topics are: student interaction, instructional management, and resource utilization. The next three items involve more personal concerns and include: motivation, student evaluation, and public relations. These relative ratings provided some general notions about how the eight topics might be ordered, but the important information from the survey is to be found in the comments provided by the beginning teachers in a section reserved for personal responses. The instructions for the open-ended section dealing with teacher comments read as follows. "Please indicate below any other activities that you would like to be included in a series of in-service help sessions for beginning teachers."

Classroom Discipline

All but eight respondents scored classroom discipline as a workshop topic of greatest importance. The survey described classroom discipline positively using these two descriptors: (a) maintaining a positive and open learning environment, and (b) motivating students. Even with this positive definition, classroom discipline was the topic of greatest concern for these beginning teachers. These respondents' comments are illustrative of the importance of discipline:

> DISCIPLINE is the biggest problem—"what to do when," when nowadays teachers aren't allowed to do much aside from give detention. (Big laugh for most students.)

> Much more emphasis must be given to the need for effective discipline. Without good discipline teacher stress "sky rockets" and the greatest lesson plans fail.

> I feel it is imperative for new teachers to have some kind of help in the area of school discipline. I see so many things now, that could have helped me then, had I known them or understood the extreme importance of little things, like not allowing students to be dismissed until you say, and beginning work immediately after the bell rings before students can get rowdy. Knowing these things in September would have made my life a lot less stressful.

Some respondents faulted their preservice training.

> "discipline" seemed like a dirty word, at least to me, when I asked questions of professors at university.

> Dealing with classroom discipline is an area that can be really improved. I cannot remember covering this at the University.

Another respondent offered a constructive way of addressing the problem.

> If I could have attended seminars with first year teachers or had more pregraduation experience in disciplinary problems I would have had less trouble. And I don't mean just student teaching. It is not a true experience in the classroom when the co-operating teacher is there with you. New teachers need practical skills perhaps even more than educational or theoretical expertise.

Student Interaction

The second most important topic identified by beginning teachers was student interaction. The examples listed for this topic were, (a) perceiving the needs and problems of students, and (b) resolving student conflicts. Among the comments that specifically addressed this topic, several respondents indicated a weakness in their preservice training.

> I feel that my training in university was very inadequate in terms of the above instructional management and student interaction. I feel that we need more time spent on the student. How he/she feels, how we can recognize problems. In order to teach a kid at any age we need to know these things at the earliest time possible.

> Basically the biggest problem I see in the teaching profession is the lack of preparation for the teacher first entering the classroom. By this I mean the first year teacher believes that this is going to be a truly enjoyable experience only, in certain circumstances, to be faced with a pack of hungry animals.

> Hands-on experience is really the only way to learn about the teaching profession, but if an in-service is absolutely necessary then concrete activities involving real problems that arise in a school such as abuse, pregnancy, etc., can be beneficial. Often the teaching is the least of teacher's worry!

Others saw a need for help for both students and teachers.

> Positive reinforcement to students and teacher—You are doing a good job.

> Effective planning of group activities to allow for more individual teaching times with one on one sessions for weaker students.

> I believe we are not prepared for the demands placed on us by the system and sometimes the lack of administrative help that we deserve. In other words, they do not have the answers to give us either because they don't know them or have the guts to explain them. Our system has failed in some respects.

Instructional Management

The instructional management topic included two examples: (a) planning and scheduling classroom activities, and (b) setting instructional goals. The rating of this item is of particular interest because these competencies are fundamental survival skills that are normally thought to be stressed in the preservice program. These beginning teachers perceived a need for help in this very practical area.

> There is a need for more practical work such as: how to fill out a register, how to assess student progress, creating tests, how to deal with students who are constantly disruptive. Classroom management, not just psychological theory.

> Arranging the curriculum would be a valuable service. In many positions this is left to the individual teacher and as a first year this is difficult.

> Some activity that would simply explain many of the mechanics or day to day operations that are never explained. Such things might include ordering, doing registers, and different regulations in terms of where teachers stand on certain issues. I feel new teachers are not well-prepared for the administrative side of teaching. This very important area should be addressed.

> A credit course with classroom management as the primary focus would be a very good idea. Preferably taught by someone with recent teaching experience.

Several teachers testified to the hard work of simply surviving the first year of teaching.

> I have so far had a very good year. However I have worked very hard and spent a huge amount of time on both instructional management (most days in school until 6 p.m. at least), and in tracking down the work of some students who do not seem to be interested no matter what I do. In particular, one child is a problem in that he has done no work in school in the past four years.

> Being on your own in a homeroom class of 33—all those faces, not knowing procedures, bells, who is allowed and not allowed to do what, methods of discipline, etc., etc. When I look back I certainly don't know how I did it, but it was due to my six years at university.

Resource Utilization

Resource utilization may be seen as a prerequisite to effective classroom management. This topic was defined by these two examples: (a) locating resource people and materials, and (b) using resources effectively. Access to resources may have posed a problem for some of the beginning teachers because of their location in rural areas, but several comments seemed to underlie the general importance of this topic.

> Throughout our professional training we were taught to vary our instructional mode through the use of audiovisual resources. These resources, however, are unavailable in many cases and we end up boring ourselves and our students to death!!

> All beginning teachers should be made aware of the materials available to them by the board. Resource utilization in the area of Special Service Student

> Program and Assessment for the High School is minimal for High School Teachers.
>
> How to use resource personnel and how to deal with the lack of resource personnel.

Personal Motivation

This topic seems to be more concerned with the personal emotions and the maintenance of a positive self-concept. It has been stated that the first four topics are more concerned with the everyday operation of the classroom. It is interesting to note that the personal concerns of beginning teachers rank below the more mechanical activities of running the classroom. The examples listed for this topic are: (a) receiving emotional support from other teachers, and (b) retaining a positive attitude in spite of stress. Several teachers clearly expressed their need for help in this area.

> I believe we are not prepared for the demands placed on us by the system and sometimes lack the administrative help that we deserve.
>
> If you have a supportive staff that alleviates a lot of possible problems for beginning teachers. I personally don't mind asking any question, and that helps because this year I had a lot of questions—lines of communication have to be kept open.

Other teachers wanted the more experienced faculty members to be tolerant in their dealings with beginning teachers.

> Beginning teachers need to be informed about many things that most teachers think we (beginning teachers) already know.
>
> Not enough emphasis can be placed on receiving support from other teachers, to know that we are all having some problems and that it will take time to become a good teacher—years in fact.

Others suggested that some form of a support mechanism be put in place.

> I perceive a need for more guidance in terms of evaluational planning and showing what should be emphasized over the year. I also believe that alternatives should be given to first year teachers in terms of discipline; what can be done.
>
> I would be especially interested in a program like (STET—Strategies for Effective Teaching) which solves strategies for real-life class situations.

Student Evaluation

Student Evaluation seems to be more closely allied to the mechanics of teaching than the personal interaction skills that immediately precede and follow this topic. The terms describing this topic were: (a) constructing and administering tests, and (b) assessing student progress. Some of the following comments indicate the importance of the topic to the beginning teachers.

> I feel a better understanding of the provincial standard of achievement would better help teachers in areas of subjective evaluation.
>
> There is a need for more practical work such as: how to assess student progress/creating tests.

> I perceive a need for more guidance in terms of valuational planning and showing what should be emphasized over the year.

Public Relations

This topic, once again, seems to involve the more personal concerns of beginning teachers and therefore fits more closely with the personal motivation topic listed above. The questionnaire referred to: (a) conducting parent conferences, and (b) maintaining a positive relationship with the community. Odell, Loughlin, and Ferraro (1987) list parent-teacher relationships as a need of beginning teachers; however, it ranked quite low among the needs of these respondents. In fact the only comment listed mimicked those identified above as examples in the survey. This suggests that these beginning teachers are quite comfortable dealing with parents and have no difficulty coping with the expectations and values of the communities where they begin their careers. This is an interesting contradiction considering the comments that were received concerning legal rights and responsibilities of teachers that are explored in the next section. Perhaps it is because students returning to rural areas with an education are awarded a certain amount of esteem, or it may be the vestiges of deference to authority traditionally shown teachers in rural areas.

Communication Skills

This topic was clearly of least concern in the survey. Over twice as many teachers listed this as of least importance than any of the other topics. The examples for the topic included: (a) making clear oral presentations, and (b) expressing written ideas clearly and concisely. Although the beginning teachers may have felt their preservice program had some deficiencies, they seemed to feel confident in their ability to make oral and written presentations to their students. There were no individual comments that pertained directly to this topic.

Serendipitous Findings

Because of the nature of beginning teacher employment in Newfoundland, three unanticipated findings emerged from the individual comments of the respondents. It is speculated that these findings may be in part a result of the difficulty that teachers have in obtaining initial employment. More specifically the three findings were: difficulties associated with teaching in a rural area, that is, few resources and multigrade classrooms; problems attaining employment in a tight job market, that is, substitute teaching and rural posting; and concerns of teaching in an increasingly litigious society, that is, rights of abused children and rights of beginning teachers.

Approximately 74% of the teachers in this study began their teaching in rural Newfoundland. It is quite apparent from conversations with graduating students that they would usually prefer to teach in the larger communities of the province, but because of the tight job market they are obliged to accept initial employment in the more remote rural areas. Beginning teachers often take a first job in a rural area in the hope that this will lead to future employment elsewhere closer to town. Some comments reflected the

difficulty of finding initial employment and the attendant problem of having to begin teaching in a rural area.

> Beginning teachers should be enlightened on how difficult it is to get a job in the province. What has been most disturbing and stressful for me has been the hiring practices of some school boards, dealing with this kind of stress is a startling realization.

> I think the majority of first year teachers are employed in smaller schools and many resources are unavailable due to lack of funds available. This issue needs to be addressed.

The other indication of a tight job market is the large number of teachers who begin their teaching career as substitutes. Often these are those who are unwilling to move to the rural areas in order to get a job. Substitute teachers are having a significant impact on teaching in Newfoundland. *Toward 2000* (1990), a recent report from the Department of Education, indicates that substitute teachers represent 22.5% of the total teaching force in Newfoundland. This substitute teaching figure represents 90,000 days or the equivalent of 473 additional full-time teachers in the system.

One beginning teacher suggested that there should be a preservice preparation program for teachers who seek to enter the profession through substitute teaching.

> For many first year teachers substitute teaching is a reality. Have a course in which students can prepare themselves for substitute teaching. Many teachers give up this profession because they find themselves in awkward situations as a substitute teacher.

The second problem resulting from beginning teaching in a rural setting may be a sense of inadequacy because of the wide range of expertise needed to deliver the curriculum. Because there are often fewer teachers in a rural school, there may be unique problems arising that were not addressed in the preservice program. As one beginning teacher put it, "we need help in knowing how to teach in a multi-grade classroom," and another was more emphatic when saying, "we need to know how to prepare for 18 courses per day."

The third unsolicited finding resulting from the survey concerned the legal rights of teachers and students. The following comments reflect this concern.

> Beginning teachers' rights and obligations. Sit down and discuss all aspects of a principal's job and where it comes down to him controlling the discipline problems. Make sure there are no misconceptions when it comes to the delegations of responsibility.

> Many beginning teachers do not have the necessary knowledge about procedures to follow in decision making about students and also teachers' rights.

Other comments reinforcing this legal concern indicate these beginning teachers wanted to know about "lay-offs, seniority and reassignments." Beginning teachers have a two-year probation period in Newfoundland, and in a tight job market it may be extremely important to know one's rights regarding dismissal.

Beginning teachers' interests in their legal rights and obligations may also be due in part to a very practical concern about sexual abuse in this society. One respondent expressed concern for knowing how to provide warmth and comfort for children and yet not being perceived to be making unwanted sexual advances.

> Some reaction sessions, how would you respond if ...? could be beneficial. These would include the teacher's legal limits as well, for example, how far can we go to help on comfort, etc.

These legal concerns were quite unexpected and probably represent the most interesting findings of the study.

Discussion

The purpose of this article is to describe the needs of beginning teachers that might be addressed as a part of induction programs. Induction tends to emphasize the cooperative role played by both teacher education institutions and local school boards in composing induction programs. Such programs are constructed to assist the employed beginning teacher to cope successfully with the transition from student into teacher. As Odell, Loughlin, and Ferraro (1987) claim,

> The questions asked by beginning teachers seem eminently sensitive to the demands placed on them by their students' needs, by the school calendar, and by parents and administrators. Through tracking the rise and fall of the teachers' several needs, the assistance offered to new teachers can be made precisely appropriate to their contemporary needs. Thus, the content of teacher induction programs can be made reactive rather than prescriptive and, to the extent that individual teacher needs are identified, it can be individualized for a particular first-year teacher. (p. 56)

Zaharias and Frew (1987) summarize a similar position this way.

> While the purpose of this article has been to demonstrate that colleges of education can make a significant contribution to the induction of beginning teachers, it has not been the intent to suggest that institutions of higher education assume total responsibility for the delivery of induction programs. Instead, the induction phase of teacher education should optimally serve as a context in which local school systems and colleges of education work together. (p. 54)

Induction in Rural Areas

Some of the most interesting findings of this study seem to have resulted from the rural positions to which the beginning teachers have been assigned. Some of the difficulties arising because of their rural posting are quite understandable. Beginning teachers in rural areas may be requested to teach a number and variety of courses for which they have little preparation. This is not due to a misassignment of beginning teachers, as may be the case in urban areas, but rather a matter of necessity. Often teachers in rural schools are expected to teach a wide variety of courses. Beginning teachers in this survey indicate that it would be helpful to have a special preservice training program for teachers who may want to teach in rural areas, with some

exposure to this problem. Beginning rural teachers may also need induction programs that differ from their urban counterparts.

In rural schools the problems of induction may be magnified and not as easily solved as in larger urban centers. As a result of this, some unique solutions must be devised. Because of the rural nature of Newfoundland and the placement of most beginning teachers in these areas, it has been proposed that beginning teachers might be assisted through a series of induction programs on the teleconference network which includes a two-way audio-visual communication capability.

This teleconferencing solution is generally applicable, and qualified persons can be gathered together in central broadcasting locations for the purpose of providing induction programs to teachers in isolated communities. Induction programs are most frequently conducted by mentor teachers, but this method of induction presents several practical problems for teachers in small rural schools. It may be undesirable to release the "best" teachers for the purpose of mentoring because there are so few that they are greatly needed in the classroom. In some cases there are no mentors available at the same grade level or even teaching the same courses. The nearest teacher with experience in the same course or at the same level may be several miles away both in distance and in spirit. One of the greatest problems for beginning teachers is having the feeling that they are all alone in this situation and that only they are experiencing these difficulties. The teleconference method would allow beginning teachers to talk with each other as well as the knowledgeable induction practioners.

Another method of offering induction programs to rural areas might involve establishing a team of itinerant teachers. These experienced people would need to have expertise in the various areas of need that might arise in the beginning teacher situations. This team would be made available to schools, school board coordinators, and/or school systems. The chief function of the team would be to offer induction programs for beginning teachers on the topics of expressed need. These induction teams might serve as trouble-shooters for schools where the mentoring systems might not be practical. If induction programs are to be offered for beginning teachers, they must be offered to all beginning teachers, both urban and rural. It should be recognized that these programs may differ significantly in both content and delivery.

Legal Concerns of Beginning Teachers

Some of the legal problems that arise in this study are perhaps unique to the Newfoundland setting. The province has recently been undergoing a prolonged investigation of child physical and sexual abuse that has been broadcast on the local cable channel each day for nearly two years. Because of this climate, teachers express fears that their overtures of affection may be misinterpreted as sexual abuse.

Newfoundland teachers are also concerned because of revisions that are being made to the Schools Act that will eliminate corporal punishment as a means of discipline. Beginning teachers in Newfoundland are not alone in their concerns about the use of discipline in the schools. Some recent court

cases have put the disciplinary role of teachers in doubt. Green (1988) reports on a case in British Columbia where a teacher was suspended for breaching a board regulation forbidding corporal punishment. In another case in Ontario, Green (1990) reports that a principal was charged with using undue force in restraining a student. Foster (1986) believes that teachers should be sued for educational malpractice, mentioning several cases where students are liable for educational harm because of having been placed in a special education class. Beginning teachers are often placed in the classes with mainstreamed students, and therefore may need to be concerned about educational negligence in this kind of a situation. Parker-Jenkins and Osborne (1985) mention several cases involving the private life styles of teachers that have left teachers in doubt about their legal status as teaching professionals.

These beginning teachers are the first generation to have had the option of taking a course in educational law as part of their preservice training. They may even be required to teach a course in law or democracy as part of the secondary school curriculum. They are much more aware of the problems and possibilities of teaching in a litigious society.

Concerns about the possibilities of layoffs, redundancies, and proper teacher conduct, especially when on prohibition, are of interest to all beginning teachers. Rural schools constantly face the problem of trying to hire and retain teachers, so the legal concerns expressed by the respondents in this survey may be quite typical for other rural areas in Canada. It is quite likely that teachers beginning their careers when there is a tight job market will experience the same legal concerns about job security that were expressed in this survey.

It is evident from the comments of the respondents in this sample that many of the difficulties experienced by beginning teachers are not fully addressed in their preservice education. The fact that these beginning teachers recognize problems and are willing to seek help to address these concerns assures us that they are committed professionals. Helping these teachers in their first years may be the key to keeping competent and highly motivated people in our education system.

References

Bullough, R.V., Knowles, G.J., & Crow, N.A. (1989). Teacher self-concept and student culture in the first year of teaching. *Teachers College Record, 91*(2), 209-233.

Cole, A., & McNay, M. (1989). Induction programs in Ontario schools: Raising questions about preservice programs and practica. *Education Canada, 29*(2), 4-9.

Darling-Hammond, L. (1989). Accountability for professional practice. *Teachers College Record, 91*(1), 59-80.

deVoss, G. (1978, 1979, 1980). *Technical reports: Institutional follow-up studies*. Columbus, OH: Ohio State University College of Education.

Focusing our future: The report of the Presidential Committee to review teacher education. (1988). St. John's, NF: Memorial University of Newfoundland.

Foster, W.F. (1986). Educational malpractice: Educate or litigate. *Canadian Journal of Education, 11*(2), 122-151.

Fullan, M., & Connelly, M. (1987). *Teacher education in Ontario: Current practice and options for the future, Position Paper*. Toronto, ON: Ontario Institute for Studies in Education.

Green, M. (1988). Case note: Employment rights. *Education and Law Journal, 1*(2), 230-231.

Green, M. (1990). Case note: Discipline. *Education and Law Journal, 2*(3), 351-353.

Griffin, G. (1985). Teacher induction research issues. *Journal of Teacher Education, 36*(1), 42-46.

Grimmett, P. (1984). *Research in teacher education: Current problems and future prospects in Canada*. Vancouver, BC: Centre for the Study of Education, University of British Columbia.

Hawk, P. (1987). Beginning teacher programs: Benefits for the experienced educator. *Action in Teacher Education, 8*(4), 59-64.

Howey, K. (1988). Mentor-teachers as inquiring professionals. *Theory into Practice, 28*(3), 209-213.

Huling-Austin, L. (1986). What can and can not reasonably be expected from teacher induction programs. *Journal of Teacher Education, 37*(1), 2-5.

Katz, L., Raths, J., Mohanty, C., Kurachi, A., & Irving, J. (1981). Follow-up studies: Are they worth the trouble? *Journal of Teacher Education, 32*(2), 18-24.

Kozioh, S.M., & Burns, P. (1986). Teachers' accuracy in self-reporting about instructional practices using a focused self-reporting inventory. *Journal of Educational Research, 79*(4), 205-210.

Loadman, W. (1983). *Technical report: Follow-up studies*. Columbus, OH: Ohio State University College of Education.

McDonald, F. (1980). *The problems of beginning teachers: A crisis in training* (Vol. I). Study of induction programs for beginning teachers. Princeton, NJ: Education Testing Service.

McEvoy, B., & Morehead, M. (1987). Teacher induction: What can a university do best? *Action in Education, 8*(4), 45-49.

McNay, M., & Cole, A. (1989). Induction programs in Ontario: Current views and directions for the future. *Education Canada, 29*(1), 9-15.

Odell, S.J. (1986). Induction support of new teachers: A functional approach. *Journal of Teacher Education, 37*(1), 26-29.

Odell, S., Loughlin, C.E., & Ferraro, D.P. (1987). Functional approach to identification of new teacher needs in an induction context. *Action in Teacher Education*, 8(4), 51-57.

Parker-Jenkins, M., & Osborne, J, (1985). Rights in conflict: The case of Margaret Caldwell. *Canadian Journal of Education, 10*(1), 66-76.

Schlechty, P. (1985). A framework for evaluating induction into teaching. *Journal of Teacher Education, 36*(1), 37-41.

Snow, S. (1988). Concerns of the first year teacher. In P. Holborn, M. Wideen, & I. Andrews (Eds.), *Becoming a teacher*. Toronto, ON: Kagan and Woo.

Toward 2000: Trends Report Two: Elementary and Secondary Projections. (1990). St. John's, NF: Division of Evaluation and Research, Department of Education, Government of Newfoundland and Labrador.

Varah, L., Theune, W., & Parker, L. (1986). Beginning teachers: Sink or swim? *Journal of Teacher Education, 37*(1), 30-34.

Wideen, M.F. & Holborn, P. (1986). Change and survival in faculties of education. *Interchange, 17*(1), 33-47.

Zaharias, J.A., & Frew, T.W. (1987). Teacher induction: An analysis of one successful program. *Action in Teacher Education, 9*(1), 49-55.

Zimpher, N. (1986). *Using research knowledge to improve teacher education: Implementation of an induction program for inquiring professionals* (Contract No. 85-0012). Washington, DC: Office of Educational Research and Improvement.

Zimpher, N.L., & Rieger, S.R. (1988). Mentoring teachers: What are the issues? *Theory Into Practice, 28*(3), 175-182.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 1, March 1991, 19-30*

MARGARET R. OLSON
University of Alberta

Supportive Growth Experiences of Beginning Teachers

This study focuses on the ways in which the experiences of support from within schools affected the professional growth and sense of satisfaction of four first-year elementary school teachers. It is based on the phenomenological analysis of interview transcripts from a larger study of the experiences of first-year teachers. The present study describes an area that was not analyzed in the original study. The themes that emerged in the present study were (a) support that allows a growth choice, (b) need for reciprocity, (c) need for appropriate levels of support, (d) need for concreteness and specificity, (e) support through comparison with others, and (f) need for recognition.

Although the need for support during the first year of teaching is well documented (Everett-Turner, 1985; Goodman, 1987; Lortie, 1975; O'Neal & Hoffman, 1984), there is a lack of research into how novice teachers experience the support that they do receive. Researchers have also shown that novice teachers get most of their career development support from peers (Goodman, 1987; Griffin & Hukill, 1983; Lacey, 1977; Lortie, 1975).

Williams (1986) points to five stages of career development in teachers: (a) survival, (b) consolidation, (c) expansion, (d) disillusionment, and (e) career stance which can be either holding or lifelong learning, and states:

> Teachers who find a way of integrating their idealism with the realities of the system seem to assume a career stance of lifelong growth. They take a vital interest in their students and their careers and continue to grow and learn. (p. 197)

If this is true it would seem essential to examine more closely the kinds of support that help foster this integration. It also seems logical in light of Ryan's (1980) findings that levels of satisfaction experienced by first-year teachers tend to continue into later teaching, to focus on the kinds of support that are effective in promoting growth and satisfaction in first-year teachers.

Williams also states that "in an atmosphere created to account for the special needs for beginning teachers, the teacher may move more quickly and easily through the first two stages of development" (p. 206). This atmosphere must be looked at more closely in order to help us understand the types of support that are essential to the professional growth of first-year teachers.

This study explores examples from the experiences of four first-year teachers in order to search for a deeper understanding of how their experi-

Margaret Olson has had extensive experience as an early childhood educator and is currently in a doctoral program with the Department of Elementary Education. Her main area of interest is qualitative research in teaching and teacher education.

ence of support from peers affected their professional growth and sense of satisfaction. Williams' question, "What is the difference between an event which causes growth and one which does not?" (p. 210) is also kept in mind.

Method

This study is based on only a small portion of the data collected for a much larger study of the experiences of first-year teachers (Olson, 1989a). A phenomenological approach was used in order to develop a deeper understanding of the often taken-for-granted problems of first-year teachers. As Fullan (1982) emphasizes, "The message to everyone outside the role under review is: *Understand the subjective world—the phenomenology—of the role incumbents as a necessary precondition for engaging in any change effort with them"* (p. 120).

Participants

The four teachers involved in this study, three women and one man ranging in age from 26 to 35 years, had just completed their first year of teaching in elementary schools in a large urban school district. Barbara, Linda, JoAnne, and David were selected because of their willingness to spend time writing a short description of their experiences, to be involved in a taped interview, and to respond to the interpretations of their experiences.

Barbara began her teaching career at age 35 after working for several years as a hairdresser. Sharing her daughter's experiences with school reintroduced her to the education system. When she started to consider a career change, her experiences with her daughter led her to consider teaching. She believed that teaching was an important, well respected profession in which she could make a contribution to society. She also felt that school hours would fit with her family responsibilities to her son and daughter. Barbara's actual placement in her first year was a part-time position teaching French to two grade 4 classes.

Linda began teaching at 33. Her previous experience as a library technician had provided her the opportunity of working in several educational settings. She had worked in both elementary and secondary schools as well as at a community college where she also instructed courses. When she decided to make a career change, her previous experience with elementary school children led to her choice. Linda's initial placement was a grade 5 classroom where she taught extended French. She taught language arts, social studies, art, and health in French as well as teaching French language. She taught her own class in the morning and then taught extended French to two grade 6 classes in the afternoon.

JoAnne was 28 years old when she began her teaching career. She had not been satisfied in her previous work experience, but did enjoy training other people. The satisfaction that she did feel in teaching and helping others led to her decision to become a teacher. Once she entered the Faculty of Education, she selected special education as her major because of her interest in the personal aspect of teaching. JoAnne's actual placement was a newly created position teaching a special education class for behavior disordered children. She was hired because of her specialized training. She had six male students in her class ranging in age from seven to 10 years.

David began teaching at age 26 after six years at university where he received an Honors BA in political science and an after-degree BEd. His experience in working with youth groups, especially Ukrainian scouting, led to his desire to become a teacher where he would have more opportunity to work with children. Although he works in Alberta and has friends here, home and family are in Ontario where David returns for holidays. David's actual placement in his first year was in a grade 4 classroom where he taught language arts, art, and French. He also taught French to two other grade 4 classes.

Procedure

The interviewer met with each of the participants individually to discuss the purpose of the study and to familiarize them with phenomenological research methods described by Becker (1986), Giorgi (1975), and van Manen (1984). They were given a copy of the research plan which asked for a short written description of the experiences that were meaningful to them as first-year teachers. The following questions (Olson, 1989a) were given as a stimulus for the written descriptions:

> How were your experiences different from what you expected? In what ways did your colleagues help or hinder you in your transition into teaching? How did the students influence the way you implemented your theoretical knowledge into actual classroom practice? How did the physical structure of your school and classroom affect your teaching style? Did some words take on different meanings than they had previously? In what ways did you feel unsure of yourself; in what ways did you develop confidence? What made you comfortable; what would you like to see changed? How did teaching affect you physically and emotionally? (p. 130)

However, it was emphasized that it was their experiences that were of interest; the suggestions were only given to stimulate ideas. Anything that they felt was relevant was exactly what was wanted. The written descriptions provided a basis for later taped interviews.

Interviews were held in the teachers' classrooms after school when interruptions were less likely. Participants were asked to expand on the themes of their written descriptions. The interviews, however, were not confined solely to these themes. Other aspects of the experiences were often brought to light. The taped interviews, ranging from 45 minutes to two hours in length, were concluded when participants felt that their experiences had been sufficiently covered.

Data Analysis

Initial written descriptions were read in order to get a feeling of what the experience had been like for each individual. Preliminary themes were identified. Similarities and differences among the four written descriptions were noted. Differences were not discarded, but looked at more closely. Interviews were held after preliminary analysis of the written descriptions.

Transcribed taped interviews were combined with the original written descriptions to comprise individual protocols which were broken down into isolated meaning units, paraphrased, and given descriptive themes. Once the themes were identified for each individual, they were grouped for that in-

dividual in a first order clustering that allowed a higher level of meaning to be abstracted. The first order clusters of themes were grouped once more into a second order cluster to reach a higher level of abstraction for each individual. The abstractions and interpretive analyses were taken back to the individuals for validation.

The themes from each of the four individual second order clusters were then combined in a final higher order cluster of common themes. Care was taken in looking at differences and discrepancies as well as tracing the reliability of the common themes back through the individual experiences. A final synthesized interpretation based on the higher order clustering of common themes was written to describe the essential qualities of the first-year teaching experience for the four participants. The synthesis was also taken back to them for validation.

Although Barbara, Linda, JoAnne, and David had very different specific experiences during their first year of teaching, as the data were analyzed 10 major common themes emerged. These were (a) orientation of the novice, (b) sense of responsibility, (c) ability to meet orientation goals, (d) taking control, (e) need for affiliation, (f) finding security—physical/emotional, (g) meeting content and process needs, (h) evaluation of goal success, (i) search for understanding, and (j) changes in role orientation.

The need to limit the length of the major study prevented the researcher from developing in-depth phenomenological interpretations of the numerous subthemes that were embedded in the 10 major common themes. The present interpretive analysis investigates one of the subthemes that was not considered in the major study: the ways in which the experiences of support connect the major themes of *finding security* and *need for affiliation*. Classroom atmosphere, a subtheme of the common theme *finding security,* has been investigated in *Room for Learning* (Olson, 1989b).

Interpretive Analysis of the Data

Support that Allows a Growth Choice

A sense of not knowing and apprehension existed within these first-year teachers as they entered a new environment in a new role, both of which felt strange and foreign. As David commented, "I was very distracted by the various goings on. I had to deal with the new environment and my new 'self'—a teacher." Bollnow (1961) states, "The space where a man finds himself at the moment may not be the space to which he belongs" (p. 32). For David, not only did he need to find a sense of belonging, but he also needed to find his new self within this strange space. The way the novices were welcomed into their schools as well as the way they perceived the school atmosphere had a tremendous impact on professional growth. When a sense of belonging was felt, they were able to begin searching for their new self as a teacher; when the space continued to feel foreign, they were unable to take the risks involved in the search for a new identity.

After the initial excitement of being offered a job, concerns about the actual placement took over. The novices needed to find a place where a sense of belonging could be felt. David was placed in the dubious position of being able to choose between two schools. Although he was glad to be in a position to choose, he had no idea of how to make the choice. He suffered a severe

sense of anxiety, fearing that he would not make the right choice. However, he states, "I think what really swayed it were the interviews themselves. Strictly based on the personal experience with the principal." David did not regret his choice. The sense of belonging that he unconsciously perceived in the interview was confirmed for him throughout his first year. "The staff at this school is absolutely wonderful. There is a general warmth that evolves in the school where one can feel secure."

Linda also felt a sense of contentment and belonging in her placement. Her previous experience in the school as a student teacher provided her with a sense of familiarity. "I had an opportunity to be in the school before and I had met some of the teachers and the principal. I liked the atmosphere here." Her initial sense of belonging led to feelings of comfort that allowed her to risk dealing with the stresses and concerns that she experienced in searching for her new self.

JoAnne was at first more concerned about the way her behavior disorder class and program would be accepted in the school than she was about her own personal acceptance. She initially saw her class and herself as a separate entity in the school. In fact, her timetable was originally different from that of the rest of the school which prevented her having contact with other teachers before school and at lunch. This was changed when JoAnne expressed a need for herself and her students to have a regular lunch break like the rest of the school. She received specific guidance from the consultant as well as perceiving administrative support and collegial acceptance for herself, her students, and her program. "Once I got to know the other teachers they were terrific—just a super staff and always ready to help me out with the materials and so on."

For Barbara the situation was very different. She had ambivalent feelings about being offered a part-time position. It was not the job she wanted, but at least it was a job. Teaching French to students who came from other classrooms led her to feel like an appendage instead of an integrated member of the staff. She did not initially have a classroom of her own. Not having a room of her own made her feel out of place, lost, and isolated. "Everybody else was so busy doing things. They were all in their classrooms and I thought 'What am I supposed to be doing?'" Being given a classroom made her feel a lot better. "This is *my* room, *my* place. That was really important." When she had a place to belong, she was able to take some risks in searching for her new self as a teacher. However, because she was teaching other teachers' students, she felt a real need for her program to be accepted by those teachers. "With one teacher it was fine. With the other one, I thought he felt that he was losing one hour out of the day. [My program] seemed to really upset him so that was kind of uncomfortable." This lack of acceptance prevented Barbara from developing a feeling of belonging. Because Barbara never did accomplish a sense of belonging with the staff, the support that she perceived was minimal. She often compromised her own needs in a desperate attempt to fit in. "It was really hard to feel your way around to see what you can do and what you can't." However, she did find her need for belonging met at inservices in her subject area where she perceived a sense of commonality and emphathy:

> Meeting with other teachers in other schools who teach the same subject area has been very helpful and therapeutic. Discussing common problems often resulted in finding solutions, but more often resulted in a tremendous release of tension. I was out of the school and I felt like I could say maybe what I felt like saying. I didn't have to be positive all the time.

These first-year teachers experienced an incredible sense of inadequacy. They had a tremendous desire to experience immediate competence and be seen as competent by others, yet had no experience and little awareness of the actual demands that their new role entailed. As David stated, "I had certain ideas about what a teacher was, but I was feeling my way around, constantly grasping at straws. Am I this image of what a teacher is?" Integrating previous self-perceptions with newly perceived expectations of oneself as a teacher requires a sense of acceptance in order to risk making the growth choices that Maslow (1986) sees as necessary for movement toward self-actualization:

> Let us think of life as a process of choices, one after another. At each point there is a progression choice and a regression choice. There may be a movement toward defense, toward safety, toward being afraid; but over on the other side, there is the growth choice. To make the growth choice instead of the fear choice a dozen times a day is to move a dozen times a day toward self-actualization. (p. 44)

For David, Linda, and JoAnne, the sense of acceptance and belonging that they experienced helped provide them with the courage to make growth choices. For Barbara, her feelings of defensiveness toward her program and her fear of not being accepted if she could not fulfill others' expectations often led her to make regression choices.

Need for Reciprocity

The word *support* has its etymological roots in the Latin word *supportare*—to carry, bring, or convey to a place. Support also means to endure or bear (Klein, 1971). The teachers in this study needed support in order to move successfully toward self-actualization as teachers. However, they were painfully aware of the burden they might be placing on those from whom they sought support.

They needed to feel accepted as professional peers before being able to seek support in areas where they felt they should already be competent. As Linda expressed, "At first I felt hesitant because I thought, 'Well, I must be expected to know all of this.'" Her original perceptions of experienced teachers as "knowing the right answers" presented a contrast barrier with her own feelings of "doubting whether I was teaching what I was supposed to be teaching all the time." Thus she felt she had nothing to give in return for the support she needed. However, Linda's original sense of familiarity and belonging allowed her to risk deeper involvement with her peers which led to a change in her perceptions:

> I felt comfortable after just getting to know people. And to know that nobody had any definite answers either as far as a lot of things go. After a while I never felt that there was a question that was not appropriate. I felt very comfortable with that.

Linda's realization of the pervasiveness of uncertainty in the teaching profession allowed her to become more accepting of her own uncertainty. She began to look inside herself for possible solutions to problems. She then began to feel that she had something to give in return for the support that she sought. This sense of reciprocity allowed her to feel less as if she was placing a burden on those from whom she sought support. "The teacher that was next door. I used to always go and bounce my problems off her. She was very receptive. She used to do the same to me after a while."

Barbara's search to fit in, belong, and be accepted by her peers was a continual concern. "I don't want to step on any toes. I don't want them to think I'm too pushy, I don't want them to think I'm too much of a coward." However, the occasional sense of reciprocity that she experienced allowed her to feel some sense of acceptance. "We did a lot of planning together. She said, 'Pick whichever topic you feel most comfortable with, one you know something about.' It was really nice." By allowing Barbara to choose an area where she felt comfortable, the other teacher was allowing Barbara to feel accepted as someone who had a degree of professional competence that was seen as mutually beneficial. It allowed Barbara to feel worthy of accepting support in areas where she still lacked competence.

Need for Appropriate Levels of Support

Magnusson, Day, and Redekopp (1988), in their proposed intervention model to help youth in transition, state that "interventions that do not address the developmental needs of youth are a costly waste of resources that may not only be ineffective, but may also be destructive" (p. 8). This statement was certainly true for Barbara. "The principal pointed [the materials] out to me, but she didn't know anything about it, how to use it. She just said, 'Here it is, Do what you want with it.'" This principal lacked an understanding of Barbara's developmental needs. She was assuming a competence that Barbara did not possess and was providing Barbara with the freedom that an experienced teacher may have appreciated, but which Barbara perceived as a devastating lack of support.

Williams (1986) states that "teachers said they often grow because of the right supervision or help at the right time" (p. 201). However, teachers may not be aware of what their needs are. They may only be able to see their needs when someone with more experience is able to point them out.

Barbara discovered her need to seek out other French teachers from a comment made by one of her peers. "One of the teachers said, 'It must be really hard for you with French because no one has anything to share,' and I thought, Oh ... yeah, it is!"

David was fortunate in finding a peer who was not only able to anticipate David's needs, but who continually did things to help David meet those needs. "I remember stepping out of my room and when I came back another teacher had given me a whole slew of materials for the first week of teaching." When David realized that this particular teacher was not only aware of his needs, but was more than willing to help, David felt comfortable in initiating requests for support.

When JoAnne was nominated for a first-year teaching award, she attributed her success to the help she received from others and thus felt

undeserving of the recognition. A comment from her principal allowed her to see her own value more clearly:

> He said to me, "If you don't take credit for things, you're just going to burn out. You have to look at things and say, 'I did this.'" That really made a lot of sense to me. I really felt a lot better after that.

At other times these first-year teachers knew exactly what they were looking for. They experienced considerable frustration when these needs were not met and elation when they found the help they needed. David was very appreciative of two staff members who were continually available to help him meet his needs. "I was very fortunate that I met two special people on the staff who last year and this year have provided me with strength, ideas to pursue my goals."

For Barbara, the need to meet the expectations of others caused a continual sense of anxiety and frustration. "I didn't know exactly what was expected of me and that really upset me. You can do as much as you want extra, but these things have to be done and I never found out what those things were." Her peers' lack of awareness of her need for guidance did not allow them to provide the support that she needed to overcome her feelings of isolation.

Linda found other teachers very helpful in providing her with specific information and resources. However, even with support from peers, she felt a deep sense of frustration at her inability to accumulate what she considered to be an adequate variety of concrete resources with which to implement the kind of program she desired. Ideas presented at inservices may have been useful to teachers in other stages of career development. For Linda they were frustrating. "I had lots of ideas. I didn't need ideas. I needed concrete things to work with."

The four novices also gained different kinds of support from teachers in different career stages. Although experienced teachers could provide a wealth of ideas, resources, and a model for emulation, their competence could also magnify the novices' sense of inadequacy. JoAnne found this experience in her relationship with her consultant. He was extremely helpful in providing JoAnne with resources and spent a lot of time discussing her concerns. However, when JoAnne made comparisons, she felt inadequate. "The consultant had a wonderful program, just excellent and I thought about that a lot and made comparisons and that was really hard on myself. He did it. I should be able to do that too."

Experienced teachers seldom provided the same type of support found in relationships with peers at the same developmental stage. Barbara felt a close bond with a teacher who had little experience. "She'd been teaching a couple of years and she knew what I was going through."

Linda also perceived a difference in the kind of support she received from experienced and novice teachers:

> I was lucky that there were three first-year teachers at our school. We supported each other a great deal. It was moral support more or less because we were all teaching very different things. We all knew that we were going through the same thing. It was different support than what you get from other teachers.

When the teachers in this study were given support appropriate to their needs, it provided them with the resources and information that enabled them to make growth choices. Inappropriate support that did not meet present needs did not promote growth and often led to regression choices.

Need for Concreteness and Specificity

Although the teachers in this study felt well prepared to teach, their abstract generalizations about teaching, limited classroom experience, and persisting student perspective toward teaching left them less well prepared than they had anticipated. When they entered the classroom as teachers, they experienced an overwhelming need for specific, concrete resources and techniques that would help them to cope with hundreds of daily practical problems. They were also unprepared for the variety of tasks that they had never considered as part of a teacher's role but now had to perform quickly and efficiently if they were ever going to have time to accomplish their major task—teaching children.

JoAnne, with help over the summer from the consultant, felt quite confident about her program goals and her ability to meet them. However, she soon discovered that a lack of resources and her unfamiliarity with system procedures thwarted her attempts at implementation:

> There is so much that is just a total mystery. Some of it is really district specific. Like the book depository and warehouse. What is a warehouse? It would be really helpful if someone explained that to us so we would know where to find stuff.

In order to set realistic expectations for self and students, these first-year teachers turned to their peers in an attempt to quickly and vicariously absorb the practical experience that they so desperately needed. Linda "discussed with other teachers what is acceptable, where to set limits. [She] found that lots of times [her] standards were a lot higher than they should have been." Linda's sense of urgency to develop the competence that she perceived in her peers often led her to quickly accept the status quo.

David also discussed expectations with peers, but he first needed to feel that they fit with his own belief system before accepting them. "Classroom discipline aspects were more talking to people and feeling what I felt good with and what I could work with with my own beliefs." David also received incredible support in the practical aspects of his actual teaching from another teacher:

> From the very first day he made special efforts to provide me with resources, models, and suggestions for teaching. He often came into the room and sat in on lessons and then he would go over them with me. He would always emphasize the positives and reaffirm my achievements. Yet at the same time he would focus in on areas that could be improved. From these talks I felt very much a part of what it meant to be a teacher.

The novices also looked to peers for examples of resource materials to use as models for developing their own materials. As David states, "You need something to fall back on. Maybe you don't always use them, but they're

there, like a security blanket. There's something there to refer to that you can try."

Linda found her peers very supportive in providing resources that they thought might be useful. Once she had concrete examples of materials, she found it much easier to create her own. "After I had a look at a unit that was all put together I thought, 'OK. I can do that.' I always had an idea that it was so much more."

These novices seemed to be able to create new materials from models of other appropriate resources. However, adaptation of inappropriate resources to meet their own specific needs was much more difficult. JoAnne often found herself in this position because she perceived the needs of her special education students as different from those in regular classes. "I always wondered, 'Well, how do I modify that?' If I could talk to someone and say, 'Well, this is what I'm doing,' it seemed to be a lot more productive." She looked to the experience of others who would be able to see the generalities that she could not identify in her specific needs. Being overwhelmed by the need for immediate specific solutions did not allow the novices time to reflect on the abstractable generalities for which their problem was only one example.

The general nature of many inservice sessions often did not meet the need for solutions to perceived particular needs. JoAnne found one-to-one support much more helpful. "The science consultant was really helpful and just kind of let me wander through the materials and pick her brains and that was really good. I found that a lot more helpful than going to sessions."

Support Through Comparison with Others

Williams (1986) states that "beginning teachers have a difficult time determining success. They gain satisfaction from the personal and academic gains their students make. Teachers who identify items against which to measure their success refer to sources of satisfaction and reward" (p. 200).

The teachers in this study found their need to compare themselves with others a double-edged sword. On the one hand it allowed them to measure their successful growth as teachers. On the other it emphasized how much further they had to go. Peers who were aware of this dichotomy could help these first-year teachers focus on their own accomplishments and help them set realistic expectations for growth. As Linda stated:

> I would see what they were doing in their classrooms, but they would say, "I've been teaching for ten years. It took me ten years to get this together so don't feel that in your first year you have to have the same kinds of things in place."

Barbara's perceived lack of acceptance and sense of isolation often caused her to see things from a me-them perspective. This tended to reinforce her sense of inadequacy and perpetuate the feeling of being unworthy of acceptance. "They come in with a smile on their face. I've been up all night. I'm so tired and still haven't accomplished nearly as much as they have." However, she did, as Williams (1986) suggests, measure her success through the accomplishments of her students. "What I really appreciated the most was

another teacher saying, 'Your kids were really great today,' or 'Your kids were doing something on the playground that was really nice.'"

Because JoAnne was the only special education teacher in the school, she was not able to identify any common elements for comparison with other teachers in the school. "There was no one to compare to and I think that was a real problem because I would only compare to some ideal in my head and that was really painful."

Need for Recognition

One of the primary goals of the teachers in this study was to prove to themselves and others that they could actually be teachers. Movement toward this goal was measured by student achievement and by how well one fit into the teaching subculture. As one is seen as a teacher by others, one begins to believe more strongly in the possibility of really becoming a teacher.

Recognition for successful accomplishments as a teacher was highly valued. As Barbara stated, "Our principal sent the occasional Happygram or thank you note for a particular job well done. This meant so much to me. I can't really express how much." This type of recognition let Barbara know that she was moving toward her goal of being seen as a teacher. However, Barbara did not receive this type of recognition often enough. "I guess the real complaint is that there is very little recognition for teachers." This lack of recognition also contributed to her sense of not being accepted into the teaching subculture of her school.

For JoAnne, recognition also presented problems. She was seen by others as very successful; however, this did not fit with her own perceptions of her competence. "I think I got a lot of recognition and I wasn't really comfortable with it. It was like, 'I don't know if I deserve this. I'm just kind of struggling through.'" The recognition put added pressure on her to be as competent as others perceived her to be. She was afraid that others would see too much and uncover the incompetence that she was attempting to hide. "I know there's a lot of things I'm doing really badly that I just want to sweep under the carpet." Also, being singled out for a first-year teaching award conflicted with her need to fit in as a regular member of the teaching subculture.

David was also sensitive to his need for appropriate recognition. He needed to be seen in his own uniqueness as a new teacher, but also did not want to be singled out from the group. "I think what was kind of nice was that there were two of us that came into the school—two new teachers, so I wasn't the only centered one."

The need to be recognized as a teacher by others needed to be carefully balanced with the need to fit into the teaching subculture in order for these first-year teachers to feel successful in achieving integration into their new role. Recognition needed to provide a sense of fitting in, of successfully doing things the way a teacher does. Being singled out made one feel different, conspicuous, and afraid.

Conclusion

Appropriate support provided the novices with nonjudgmental acceptance into the teaching subculture. It allowed them to feel like worthy contributors who were welcome additions deserving of guidance and support. It provided

specific information that enabled them to shape individual frameworks on which to build personal pedagogical competence. It recognized and encouraged the professional growth of the novices in moving toward pedagogical ideals. It provided explicit examples which brought to conscious awareness the pedagogically implicit knowledge that they possessed but were unable to effectively implement.

Appropriate support did more than provide models of behaviors and procedures that allowed the novices to be seen as teachers. Appropriate support gave them the courage to make growth choices that promoted movement toward the self-actualization of individual potential. As David summed up, "I am beginning to feel like a teacher. The strength that I get from my colleagues has made it much easier for me to accept my responsibilities as a teacher."

With this type of support, these novices could continue to develop professionally and take pride in *being* teachers.

This particular interpretation of the data illuminates one perspective of the support experiences of four first-year teachers. Other interpretations and perspectives are indeed possible and should be encouraged in order to uncover more of the mystery of becoming a teacher.

References

Becker, C.S. (1986). Interviewing in human science research. *Methods, 1,* 101-124.

Bollnow, O.F. (1961). Lived-space. *Philosophy Today, 5,* 31-39.

Everett-Turner, L. (1985). Towards understanding the lived world of three beginning teachers of young children. *Alberta Journal of Educational Research, 31,* 306-320.

Fullan, M. (1982). *The meaning of educational change.* Toronto, ON: OISE Press.

Giorgi, A. (1975). Convergence and divergence of qualitative and quantitative methods in psychology. In A. Giorgi, C. Fisher, & E. Murray (Eds.), *Duquesne studies in phenomenological psychology* (Vol. 2, pp. 72-81). Pittsburgh, PA: Duquesne University Press.

Goodman, J. (1987). Factors in becoming a proactive elementary school teacher: A preliminary study of selected novices. *Journal of Education for Teaching, 13,* 207-229.

Griffin, G.A., & Hukill, H. (Eds.). (1983). *First years of teaching: What are the pertinent issues?* (R&DCTE Report No. 9051). Austin, TX: University of Texas at Austin.

Klein, E. (1971). *Klein's comprehensive etymological dictionary of the English language.* New York: Elsevier.

Lacey, C. (1977). *The socialization of teachers.* London: Methuen.

Lortie, D.C. (1975). *Schoolteacher: A sociological study.* Chicago, IL: University of Chicago Press.

Magnusson, K., Day, B., & Redekopp, D. (1988). *Skills are not enough: A concept paper on innovative strategies and services for youth in transition.* Unpublished manuscript, Canadian Guidance and Counselling Foundation, Ottawa.

Maslow, A.H. (1986). *The farther reaches of human nature.* New York: Penguin.

Olson, M.R. (1989a). *Learning to teach: The lived experiences of four first year teachers.* Unpublished master's thesis, University of Alberta.

Olson, M.R. (1989b). Room for learning. *Phenomenology + Pedagogy, 7,* 173-184.

O'Neal, S.F., & Hoffman, J.V. (1984). *Curriculum decision making and the beginning teacher* (R&DCTE Report No. 9058). Austin, TX: University of Texas at Austin. (ERIC Document Reproduction Service No. ED 251 439)

Ryan, K. (1980). *Biting the apple: Accounts of first year teachers.* New York: Longman.

van Manen, M. (1984). Practicing phenomenological writing. *Phenomenology + Pedagogy, 2,* 36-69.

Williams, L. (1986). *Developmental patterns of teaching careers.* Unpublished doctoral dissertation, University of Alberta.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 1, March 1991, 31-48*

CLAIRE F. STAAB
University of British Columbia

Teachers' Practices With Regard to Oral Language[1]

The study investigated how elementary teachers structure their classroom and the school day to permit students to talk. Detailed observations over a week were made in 12 randomly selected grade 3 classrooms and 12 randomly selected grade 6 classrooms to determine the amount of time spent under the conditions of Teacher Talking, Students Speaking Formally, Students Speaking Informally, and Quiet over all curriculum areas. Results indicated that most of the school day (78% of the activity time) was spent in either quiet work time or the teacher talking. Activities such as brainstorming, oral problem solving, student-led discussions and role playing received minimal time (in combination, less than 1% of the instructional activity time). Students Speaking Informally occurred most often during art, physical education, and writing periods. Female teachers allowed significantly more Students Speaking Formally than did male teachers ($F=4.97$, $p=.036$), *but neither grade level taught nor numbers of years of teaching experience was significant with regard to any talk structure.*

During the late 1960s and early 1970s, several studies indicated that talk in elementary school classrooms was dominated by the teacher. Teachers talked and students listened. Partly as a reaction to this teacher domination, and partly as a result of increasing attention being focused on the relationship of talk to learning (Britton, 1970; Langer, 1960; Vygotsky, 1962), the professional literature of the 1970s and the 1980s abounded with articles suggesting that students be given increased time during the school day for talking and that this talking be done during many different types of activities. The purpose of this study was to investigate whether there has been a shift in this teacher domination of classroom talk during the past 20 years and to investigate during which activities, if any, teachers encourage students to talk.

Furst and Amidon (1967) found that teachers talked during 45-50% of all class time, students talked during 27-30% of all class time, and 15-20% of the time the classroom was silent. Bellack, Kliebard, Hyman, and Smith (1966) found that 15 teachers talked 50% more than 345 pupils and that three fifths of all teacher talk was for the purpose of soliciting a student response. Flanders (1970), in analyzing teachers' classroom behavior, found approximately the same ratio of teacher to student talk as did Furst and Amidon. Several researchers (Edwards & Furlong, 1978; Hennings, 1975) replicated the general findings of Flanders during the 1970s. None of these studies indicated during which specific activities students were encouraged to talk,

Claire Staab is an associate professor in the Department of Language Education. She teaches classes in oral language development and has been an active researcher in the area of oral language in elementary school classrooms.

but clearly past research has revealed that teacher talk dominated the school day.

In the minds of many theorists and educators, this teacher domination of classroom talk was unfortunate as it seemed inconsistent with the growing knowledge of how children learn (Vygotsky, 1962). The idea that talk is central to the process of learning has gained prominence during the past 15 years (Chorny, 1981; Jones, 1988). The literature of the 1970s, 1980s, and early 1990s evidenced an abundance of articles suggesting that if teachers want children to take an active part in learning, it is important for the child to talk (Barnes, 1976, 1990; Barnes & Todd, 1977; Beebe, Steven, & Masterson, 1982; Edwards & Mercer, 1987). Students need time to talk and for this to happen, teachers must talk less.

Before proceeding, it seems important to note that not all research done in the 1970s and 1980s advocated the importance of student talk in the classroom. Researchers on teacher effectiveness have found that engaged time on task, academic emphasis in teaching content, and direct instruction carried out by the teacher were included among the correlates of effective teaching as measured by student achievement on standardized tests (Brophy & Evertson, 1974, 1976; Brophy & Good, 1986; Evertson, Anderson, & Brophy, 1978; Evertson, Emmer, & Brophy, 1980; Fisher et al., 1978; McDonald & Elias, 1976; Rosenshine, 1976; Soar & Soar, 1972; Tikunoff, Berliner, & Rist, 1975). Subsets of these studies conducted during reading and math lessons have suggested that students who were academically engaged under direct teacher supervision scored higher on standardized measures (Anderson, Evertson, & Brophy, 1979; Good & Grouws, 1977, 1979; Stallings, Cory, Fairweather, & Needels, 1977, 1978). Several studies (Brophy & Evertson, 1976; Soar, 1977; Solomon & Kendall, 1979; Stallings et al., 1977, 1978; Stallings & Kaskowitz, 1974) and have shown that successful teachers were strong, businesslike leaders who had high academic expectations for students. Less successful teachers were more child centered in that they allowed students to select their own activities and working arrangements (Good & Beckerman, 1978; Soar, 1977; Solomon & Kendall, 1979; Stallings & Kaskowitz, 1974; Stallings, Needels, & Strayrook, 1979). These child centered teachers often allowed students to talk with peers.

Clearly inconsistencies exist between a variety of teacher effectiveness studies and studies in language learning. Resolving these inconsistencies is impossible without further research. From the data presented in the above teacher effectiveness studies, it is difficult to ascertain the intent or amount of teacher organization present in the classrooms of less successful teachers. Did the teachers structure tasks that were to be completed by groups or pairs of individuals, or was the "talk time" simply time off task? In the Stallings and Kaskowitz (1974) and the Soar (1977) studies, nonacademic activities were defined as dramatic play, games, talk about home and family—not structured discussion groups or students helping one another on academic tasks. What parameters, if any, were placed on student self selection? Were the students trained in any way to be self selecting or were they left entirely to their own devices? These questions need to be answered. Certainly studies in which a child centered teacher was highly successful have been reported

(Edelsky, Draper, & Smith, 1983; Graves, 1983; Gunderson & Shapiro, 1986; Hansen, 1987; and Platt, 1984), and the social and academic contexts of these types of classrooms need to be compared with the social and academic contexts in the types of classrooms included in the teacher effectiveness studies.

Another important issue is the length of the data collection period. In general, observation during most of the teacher effectiveness studies took place over a period of a few days or weeks. Data from cooperative learning studies have suggested that to be successful groups must be well organized by the teacher, students must learn to work effectively in these groups, and that this learning takes time (Logan, 1986; Sharan & Shachar, 1988; Slavin, 1983). Groups of this nature need to operate for several months before a realistic evaluation is possible (Damon & Phelps, 1989; Wasserman, 1989).

Another possible reason for the discrepancies lies in the evaluation instrument. Standardized tests may not be the best means of evaluation when testing the benefits of talk in the classroom. Research has indicated that peer instruction and cooperative learning groups have led students to higher level cognitive responses than they were able to produce individually (Gilbert & Pope, 1986; Graybeal & Stodolsky, 1985; Sharan & Shachar, 1988; Trudge & Caruso, 1988). Damon and Phelps (1989), in working with 4th and 5th graders on problems in mathematics, found that students achieved deeper insights into concepts when allowed to collaborate with peers, but that peer collaboration was less successful for those tasks that relied on formulas and procedures. It is difficult to measure higher level cognitive responses on standardized tests. Phillips (1988) has suggested that the best measure of assessment is not a standardized test but the language of students in terms of their ability to put a point across, produce supporting facts, and sustain an argument.

Regardless of the inconsistencies that need to be resolved by further research, studies do exist that link student talk to learning. King (1989), in working with grade 4 students of high and average ability on computer assisted problem solving tasks, found that students working in small groups did use language to clarify meaning and to try a number of strategies for solving the problem. Webb and Kenderski (1984) reported that students' providing explanations for other students was associated with achievement, and current research is indicating that intermediate students can improve their reading comprehension through literature response groups (Eeds & Wells, 1989; Strickland, Dillon, Funkhouser, Glick, & Rogers, 1989).

Assuming, then, that talk can be a factor in the process of learning, several researchers, theorists, and educators have attempted to further explain the relationship between talking and learning. Edwards and Westgate (1987) have stated that talk allows the speaker to investigate how relationships and meanings are organized. Barnes (1990), Lindfors (1987), and Staab (1990) have noted that talk plays an important part in learning in that it helps one focus information, clarify ideas, and formulate opinions.

It is not only talk in general that is important but specifically talk with peers or talk in which the teacher acts as a facilitator (e.g., brainstorming sessions) rather than as a leader in absolute control. Chang and Wells (1988)

have noted that in collaborative talk the learner must have ownership of the task. When a teacher is in control of a group, the freedom of pupils to introduce their own ideas is highly curtailed (Edwards & Mercer, 1987). The teacher retains control over what is said and done, what decisions are reached, and what interpretations are put on experiences. Often in this situation students may not have an opportunity to clarify their ideas or to formulate their own opinions.

Recent research has indicated that students are able to formulate opinions and to clarify ideas in situations where the teacher is not in control. Eeds and Wells (1989) reported that grade 6 students were able to engage in active inquiry in literature response groups in which the teacher was only a participant. Webb, Ender, and Lewis (1986) found that in debugging computer problems, groups that carried out abstract planning strategies without the teacher were more likely to learn the material.

The benefits of peer interaction are numerous. Talk with peers often clarifies both ideas and instructions (Cohen, 1989). Exploratory talk of students in small groups (i.e., peer talk) is one means by which assimilation and accommodation of new knowledge to the old is carried out. Cazden (1986) has stated that students are given more opportunity during talk with peers to ask questions of an informational nature and that questions asked of teachers are often procedural rather than informational. The asking of informational questions is critical to the learning process.

Talk is important not only for learning but as a way to improve oral communication skills. Children should be offered many opportunities to talk, and it should not be assumed that they enter the classroom being able to communicate effectively (Berliner & Casanova, 1988; Pigford, 1988; Stewig, 1988). To improve oral communication skills, students should be offered opportunities to talk to each other as well as to the teacher (Pigford, 1988; Stewig, 1988). Peer interaction provides important social experiences that foster skill development not achieved during interaction with adults (Cryan, 1984).

Given that the past 10 to 15 years have produced an explosion of information on why student talk is important in the classroom and that this information has, in many cases, reached the form of teacher manuals (Inner London Education Authority, 1984; Province of British Columbia Ministry of Education, 1989), one might wonder to what extent this information has changed the pattern of teacher domination of talk in the classroom. No recent studies have been reported, but several educators have expressed opinions. Gambell (1980) has stated that students today are probably being given more opportunities than ever to talk, to interact orally with their peers and to learn through oral communication. Edwards and Westgate (1987) have noted that innumerable teachers are now becoming interested in oral language in the classroom and are enrolling in inservice courses that are increasing their awareness of the place of oral language in the school day. Awareness of the value of oral language does not necessarily translate into increased student talk, however.

Not all educators agree that change is talking place at the level of classroom interaction. Jones (1988) has stated that in spite of the quantity of

theoretical work in the area, there has been very little real change in the extent of productive talk that students actually engage in during the teaching day. This opinion was based on informal classroom observation rather than research. Chilcoat (1987) has agreed that little has changed since the studies conducted in the late '60s, but again this opinion is not based on recent research.

Research is needed to determine if the information on the value of oral language has reached the classroom at the level of day-to-day interaction. If change has occurred, is it more evident among any particular group of teachers? Allen, Brown, and Yatvin (1986) have suggested that if a change in the ratio of teacher to student talk has been made, it is more evident in primary classrooms than in intermediate classrooms. Edwards and Westgate (1987) have suggested that change may be more evident in the classrooms of teachers who have had advanced training (e.g., inservice) in the area of oral language. The following study is one validation of the opinions that have been expressed in the literature over the past several years.

Specifically the researcher asked the following questions:

1. How do teachers structure the school day with respect to quiet time, student talk, and teacher talk across various curriculum areas?
2. Does the way a teacher structures the school day with respect to quiet time, student talk, and teacher talk differ on the basis of grade level taught, sex of the teacher, number of years of teaching experience, or the number of reading and/or language arts courses taken?

Method

Subjects

The subjects in the study were 12 randomly selected grade 3 and 12 randomly selected grade 6 teachers from a pool of over 150 teachers at each grade level. The subjects taught in a large suburban area of Vancouver, British Columbia, Canada. In selecting the subjects for the study, the researcher obtained a list of all grade 3 and all grade 6 teachers from the school district office. Grades 3 and 6 were chosen because the researcher wanted to obtain a representation from both a primary and an intermediate grade level.

Procedure

The randomly selected subjects were telephoned at their schools and asked to participate in a classroom interaction study. At no time was oral language mentioned. Subjects were told that they would be given the details of the study when it was finished, but to preserve the integrity of the study no specific details could be provided before the data were collected. Subjects were told that a research assistant would observe in their classrooms for a period of one week and code the classroom interaction.

Classroom observations took place during the months of October and November, excluding the week before Halloween. All classroom observations in each classroom were for a period of one week and included all classroom activities from the beginning of school Monday morning until the close of school Friday afternoon. Periods before school, recess, and lunch time were not observed. Each research assistant observed in one classroom for a period of one week; thus it took four research assistants a period of six weeks to

observe in the 24 classrooms. On average each classroom was observed for about 1,300 minutes during the week.

On the Friday following each classroom observation, teachers were asked by the research assistant to provide some demographic information. Teachers were asked how many years they had taught and how many reading and/or language arts courses had been taken at the university level. The research assistant also noted the sex of the teacher and the grade level taught.

Prior to data collection, the four research assistants received one week's training on the use of the classroom observation form which appears as Table 1. Training took the form of initial explanation of the form, the coding of videotapes of classrooms, and the coding of actual sample classrooms. No sample classrooms used for training were part of the randomly selected sample used in the study.

The classroom observation form appearing as Table 1 was designed to record four basic classroom conditions: Teacher Talking, Students Talking Formally, Students Talking Informally, and Quiet Classroom. The category of Teacher Talking included four subcategories: the teacher reading aloud, the teacher lecturing, the teacher leading a question-and-answer session, and the teacher talking to individual students as he or she was helping them with their work. If the teacher spoke continuously for more than 30 seconds, the talk was coded as teacher lecturing. If the teacher asked a question of students within that 30 seconds, the talk was coded as question-and-answer even though the teacher was doing the majority of the questioning.

The category of Students Speaking Formally included those times when the teacher specifically called on one or more students to give a presentation to the class or a small group. The teacher expressly had given the floor to one student or a group of students. The subcategories for Students Talking Formally were an individual student giving an oral report, group choral reading or speaking, a dramatic production, an individual student reading aloud to a group, or a student debate.

The category of Students Speaking Informally included those times when students had the teacher's permission to speak but the talk was of a less formal, spontaneous nature. The teacher was allowing talk, but had not given the floor expressly to one student or group of students. The subcategories for Students Talking Informally were a student-led discussion either in small or large groups (the teacher may have been a participant but was not the designated leader, or the teacher was not a direct member of the group), role playing or spontaneous drama, brainstorming, group problem solving, students informally helping each other such as in a partner situation, and student initiated questions.

The category of Quiet Times was recorded when there was no oral language in the room and when the students were talking without the teacher's permission. If students were talking informally during a work time and the teacher either gave them permission to do so or made no effort to stop the talking, the time was recorded as SIHX (students informally helping one another). If students were talking during work time and the teacher asked

Table 1
Instruction Page of the Classroom Observation Form

General Instructions for Structured Observation Form

General Rules

1. Only oral language will be observed.
2. Focus of the observation is on the teacher or on the activity that the teacher assigned, not on what the students are doing if these differ.
3. The unit of observation is an entire activity (e.g., the morning exercises or the brainstorming session).
4. Each activity will be timed to the nearest minute. If an activity lasts less than 30 seconds, it will not be recorded.

Abbreviations or codes

Talk Structure

Note who is talking and in parentheses the nature of that talk. TT = Teacher talking, STF = formal student talk, STI = informal student talk, QC = quiet classroom.

TTRA	=	teacher reading aloud
TTLX	=	lecture
TTQA	=	question/answer/comment
TTSX	=	teacher helping individual student
SFRT	=	individual student giving oral report
SFCS	=	group choral reading or speaking
SFDR	=	dramatic production
SFRX	=	individual student reading aloud to group
SFDT	=	student debate
SIDX	=	student led discussion
SIRP	=	role playing or spontaneous drama
SIBX	=	brainstorming
SIPS	=	group problem solving
SIHX	=	students informally helping each other, reading to each other or talking with permission
SIQX	=	student initiated questions
QCQX	=	no oral language – quiet work time
QCUS	=	quiet work time, students talking without permission

Curriculum Activity

ME	=	morning exercises	AX	=	art
TX	=	transition	FT	=	free time
RX	=	reading	HX	=	housekeeping activities
LO	=	language arts oral	OL	=	oral language instruction
LW	=	language arts written	ST	=	show and tell or news
MX	=	math	PE	=	P.E.
SS	=	social studies	MU	=	music
SX	=	science	FR	=	French

Note: < indicates that two or more activities are happening simultaneously.

them to be quiet, or indicated that quiet was required, then the time was scored as QCUS (quiet work time, students talking without permission).

These four talk conditions (i.e., Teacher Talking, Students Talking Formally, Students Talking Informally, and Quiet) were measured across all curriculum areas. The codes for these curriculum areas appear at the bottom of Table 1. Four general rules for the observations are also shown in Table 1.

The second page of the classroom observation form, which is shown in Table 2, gives an example of how the coding was actually recorded. The coder indicated the talk structure (e.g., TTQA-teacher question/answer/comment) in the first column, the curriculum area (e.g., MX-mathematics) in the second column, the time the activity began in the third column, and the time the activity stopped in the fourth column.

If two talk structures were happening simultaneously in the classroom (e.g., one group was working quietly while the teacher was lecturing to a second group), both activities were coded and these simultaneous talk structures were indicated by a < mark.

The researcher spent one day per week with each of the research assistants to independently code the oral language in the classroom. This procedure was done for the purpose of providing a measure of interrater reliability. Two measures of time were taken. The total amount of time spent under one talk structure in one curriculum area was calculated and labeled T time (i.e., total time). The number of times a talk structure occurred under each curriculum area was calculated and labeled N time (number of incidents). N time was the total number of times a talk structure occurred without regard for its duration, and T time was the duration measured to the nearest minute. Pearson Product Moment Correlations were calculated for both T time and N time as the coding of the research assistant was correlated with the independent coding of the researcher. The correlation between the two codings for T time was equal to .8252 and the correlation between the two codings for N time was equal to .9030.

Results

Table 3 indicates the total number of minutes spent in each talk structure across all curriculum areas. SFDR (student dramatic production) and SFDT (student debate) have not been included in the table or subsequent tables as no evidence of these talk structures was observed. The number in each row indicates the total number of minutes that each individual teacher spent under the condition of each talk structure. The categories are not mutually exclusive, because if two or more talk structures were occurring simultaneously all talk structures were recorded. The total shown in the right-hand column, Total Activity Time, is the sum of the values in the row. It is an indication of the gross amount of activity in the class.

Table 4 indicates the proportion of total activity time each teacher spent in each of the major talk structures. The values were determined by cumulating the categories from Table 3 and dividing by total activity time.

Table 5 indicates the total number of minutes spent in each of the four major categories of talk structures (i.e., Teacher Talking, Students Speaking Formally, Students Speaking Informally, and Quiet) categorized by each

Table 2
Mock Sample of a Coded Observation Form

Teacher's Name_____ Grade____ School_____ Date____

Time Observation Began_____ Time Ended_____ Recess_____ Lunch_____

Talk Structure	Curric. Area	Time Started	Time Ended
TTQA	MX	10:42	10:43
TTLX	MX	10:43	10:44
QCQX	MX	10:44	10:45
QCQX TTSX	MX	10:45	10:46
QCQX TTQA	MX	10:46	10:47
QCQX TTSX	MX	10:47	11:08
TTQA	OL	11:08	11:10
QCQX	MX	11:10	11:31
QCUS	TX	11:31	11:32
TTLX	SS	11:32	11:33
TTQA	SS	11:33	11:34
TTLX	SS	11:34	11:35
TTSX QCQX	SS	11:35	11:59
TTLX	TX	11:59	12:00

Table 3
Duration in Minutes per Teacher of Talk Structures Over One Week

Teacher	QCOX	QCUS	SFRT	SFCS	SFRX	SIDX	SIRP	SIBX	SIPS	SIXH	SIQX	TTRA	TTLX	TTQA	TTSX	Total*
1	796	221	5	36	9	0	0	0	0	244	24	19	293	250	472	2369
2	284	467	11	99	9	0	0	9	1	376	0	58	259	248	490	2311
3	94	68	93	43	34	0	0	2	1	760	0	24	93	146	880	2238
4	287	186	91	0	22	31	0	7	0	469	72	12	160	202	676	2215
5	620	78	45	10	48	0	0	1	0	211	0	55	49	389	614	2120
6	489	100	37	29	59	0	0	3	0	276	4	29	96	437	533	2092
7	712	87	0	0	16	1	0	0	0	159	7	96	289	210	498	2075
8	750	37	40	8	16	0	0	0	0	173	19	31	302	170	501	2047
9	339	143	94	15	15	0	0	34	0	366	0	54	198	114	654	2026
10	166	16	0	125	31	70	0	0	9	533	100	16	190	83	555	1894
11	435	75	46	30	80	0	0	5	0	106	0	35	531	126	415	1884
12	685	46	44	5	0	100	4	0	111	97	4	0	71	175	515	1857
13	362	39	32	63	18	7	10	24	0	311	0	78	339	223	342	1848
14	546	44	56	0	2	0	0	7	0	251	0	135	244	176	346	1807
15	340	80	82	15	26	0	0	0	2	267	13	53	202	265	498	1751
16	194	85	63	62	0	12	0	0	0	515	21	0	209	252	321	1734
17	786	31	0	0	3	0	0	0	0	241	18	6	206	116	286	1693
18	600	8	104	22	14	14	0	0	3	206	2	26	164	242	274	1679
19	312	26	0	15	64	10	0	0	0	254	44	81	472	181	194	1653
20	344	60	0	109	31	0	0	0	0	317	7	30	175	266	302	1641
21	591	39	46	0	37	9	19	0	0	94	2	28	224	186	360	1635
22	528	8	13	41	109	0	0	4	0	125	5	84	111	194	368	1590
23	441	76	23	14	10	10	0	0	1	42	11	51	193	288	386	1546
24	437	12	62	5	17	0	0	0	0	317	15	50	190	113	324	1542
Totals:	11,138	2,032	987	746	670	254	33	96	128	6710	368	1,051	5,260	5,052	10,714	

QCQX --quiet, no oral language
QCUS --students talking without permission
SFRT --student giving oral report
SFCS --group choral reading/speaking
SIDX --student led discussion
SFRX --student reading to group
SIBX --brainstorming
SIPS --group problem solving
SIRP --role playing
SIQX --student initiated questions
TTRA --teacher reading aloud
SIXH --students helping one another
TTQA --teacher question/answer
TTSX --teacher helping individual student
TTLX --teacher lecutring

*NOTE: Total = total number of activity minutes (activity time)
Categories of talk structures are not mutually exclusive

curriculum area. The final column in this table indicates the proportion of time spent in each curriculum area over the total number of minutes all subjects were observed.

Tables 3, 4, and 5 present a picture of how teachers structured their school day but do not answer the second research question. The second research question explored the notion of demographics. Does the way a teacher structures the school day with regard to Teacher Talk, Students Talking Formally, Students Talking Informally, and Quiet differ on the basis of grade taught, number of years of teaching experience, number of reading or language arts courses taken at the university level, or sex of the teacher? For the purpose of analysis, the sex of the teacher and whether the teacher taught grade 3 or grade 6 was obvious. Years of teaching experience was divided into two groups, those teachers who had taught 10 years or less and those teachers who had taught more than 10 years. Number of reading and/or language arts courses taken was divided into two groups—five or less and more than five. The data were subjected to a one-way analysis of variance for both T time and N time across each of the above demographic variables across each of the four major talk structures (i.e., Teacher Talking, Students Talking Formally, Students Talking Informally, and Quiet). For the purpose of these analyses,

Table 4
Proportion of Total Activity Time for Each Major Talk Structure

I.D. From Table 3	QC	SF	SI	TT	Grade Level Taught
2	32	4	16	46	3
3	7	8	34	51	3
5	33	4	10	52	3
6	28	6	13	52	3
9	24	7	18	51	3
10	10	9	37	44	3
11	27	8	6	59	3
13	22	6	19	53	3
15	24	7	16	53	3
19	21	5	19	56	3
23	34	3	5	59	3
24	29	5	22	43	3
1	43	2	11	44	6
4	21	5	25	47	6
7	38	1	8	53	6
8	39	3	9	49	6
12	39	2	16	41	6
14	32	3	14	50	6
16	16	8	32	46	6
17	48	0	15	36	6
18	36	8	13	45	6
20	25	9	19	47	6
21	38	5	8	49	6
22	34	11	8	47	6

Table 5
Amount of Activity by All Teachers on Each Talk Structure for Each Curriculum Area

Curriculum Area	OC	SF	SI	TT	Total	Proportion of Total Activity
Art	320	0	1,026	1,371	2,717	6
Computer Instruction	5	0	69	60	134	0
French	41	29	5	147	222	0
Free Time	225	3	858	602	1,688	4
Housekeeping	808	74	711	2,484	4,077	9
Library	176	0	157	188	521	1
Language Arts, Oral	74	220	96	370	760	2
Language Arts, Written	2,952	182	1,041	4,744	8,919	20
Morning Exercises	102	56	186	481	825	2
Music	237	432	50	340	1,059	2
Math	2,402	74	845	4002	7,323	16
Oral Language Instruction	197	0	23	337	557	1
Physical Education	218	16	905	846	1,985	4
Reading	3,085	501	291	2,727	6,604	15
Socials	947	224	539	1,305	3,015	7
Show-tell	16	530	30	115	691	2
Science	470	62	259	1,117	1,908	4
Transition	895	0	498	841	2,234	5
Totals:	13,170	2,403	7,589	22,077	45,239	100

Codes:

QC—quiet classroom
SF—students talking formally
SI—students talking informally
TT—teacher talking

all curriculum areas were combined. The only significant difference at the .05 level appeared in T time for sex of teacher during Students Talking Formally. Female teachers had significantly more Students Talking Formally than did male teachers ($F=4.97$, $p=.036$). A significant difference was not found in any other analysis of variance.

Discussion

The total activity of time varied from teacher to teacher because several activities could occur simultaneously. On average, although there is a variation of 7% to 48%, teachers spent 29% of the activity time during the school day asking students to work quietly and independently. Another 49% of the activity time was consumed by teachers talking. These percentages were similar to those found by Furst and Amidon (1967). When students were allowed to talk, this talk was evidenced mostly by students helping one

another informally during work times. During these work times, typically, students were allowed to chat with one another, to ask each other questions regarding their work, or to share what they were doing. Earlier studies did not describe student talk, so a comparison of the activities that elicited student talk is impossible. This informal talk consumed 15% of the activity time. If these figures for Teacher Talk, Quiet Time, and Informal Student Talk during work time are totalled, they account for 93% of the activity time during the school day.

These figures leave 7% of the activity time to be accounted for by other forms of talk. Students spent very little time reading aloud to one another or in small groups. Reading aloud accounted for 1% of the activity time and 12 of the 24 teachers spent less than 20 minutes in a week at this talk structure. Students talking in front of a large or small group such as in Show-and-Tell, or reporting the news or a current event, occurred slightly more often—2.2% of the total activity time. Eight of the 24 teachers spent less than 20 minutes per week at this talk structure. The talk structures of student-led discussions (e.g., small group discussions in which the teacher was not conducting a question-and-answer session), role playing, brainstorming and problem solving, in combination, comprised about 1% of the activity time. Where these types of talk structures did occur, they were localized largely in one to two classrooms and absent in other classrooms (e.g., one teacher spent a total of 111 minutes in one week in a problem solving activity whereas the other 23 teachers spent a combined total of 17 minutes.

If one simply looks at the figures, one must conclude that the overall amount of students' oral language in the classroom has not changed significantly in the past 20 years. Even though numerous articles have been published and recommendations have been made that children be encouraged to engage in many types of oral language activities as a means for learning, it is apparent that these recommendations have not become a reality in today's classrooms. Corson (1984) speculated that talking takes time, and upper elementary grade teachers are too busy covering curriculum to allow time for discussion and generating ideas by means of talk. According to the present study this also is true in the primary grades.

The study did provide information as to the type of activity students engaged in when talk was encouraged. This information was not available in previous studies. Even if this information is now available, it is difficult, if not impossible, to state how much time during the school week teachers should spend on various types of talking activities. In effect, it is difficult to answer the question of how much is enough. It would be unwarranted to state that teachers should spend 10% of the school day on brainstorming, another 10% on problem solving, and so forth. Neither should teachers feel that these talk structures should be practiced for themselves as isolated activities. If students are to use oral language as a means for learning, then teachers should consider many types of talk structures as integral to curriculum areas (Staab, 1990). In mathematics children may need to brainstorm as a means of thinking through possible methods for solving a problem. In science children may need to problem solve as a way to discover alternative means for completing an experiment. In reading children may need to role play as a

means for understanding a character's motivation. In social studies children may need to form small discussion groups as a means to review what they have learned. If any tie-in can be made with the teacher effectiveness studies, time spent talking must be time on task. To increase student talk in the classroom teachers need to vary the pattern of lecturing or leading a question-and-answer session and then assigning seat work. If students are to use oral language as a means for learning, then they need many opportunities to talk about what they are doing (Beebe, Steven, & Masterson, 1982; Edwards & Mercer, 1987; Lindfors, 1987).

It is also important to note that in the study there was considerable variation among individual teachers. Previous studies simply quoted averages. It may be misleading to state what "all" teachers do. All teachers asked for quiet work time, allowed students to talk within work times, lectured, asked questions, and helped individual students, yet the range of time spent during each of these talk structures was large. Individual variation was not connected to grade level taught, number of years of teaching experience, the number of reading and/or language arts courses taken at the university level, or sex of the teacher. Female teachers did provide significantly more opportunities for students to speak formally than did male teachers, yet beyond this one difference individual teachers seemed to vary in the use of talk structures in accordance with their own teaching style and not in accordance with the above demographics. Even though it is true that few teachers engaged in such talk structures as brainstorming and role playing, there was much individual variation in the use of the more common talk structures (e.g., teacher talking, quiet time). In discussing their individual profiles with the teachers involved in the study, many seemed surprised at the lack of various talk structures that had been observed.

In interviews only 21% of the subjects felt that a quiet classroom is a more productive classroom, and 98.5% felt that oral language should be a way to learn in all content areas. Even teachers who favored a quiet classroom felt that productive talk (i.e., talk for the purpose of learning) was valuable. Perhaps teaching is such a complex act that teachers often are unaware of the amount of time they spend talking and the amount of time they allow students to talk. The reasons for this discrepancy between belief and action need to be explored in further research.

In examining the number of minutes spent in each of the four main categories of talk structures (i.e., Teacher Talking, Students Speaking Formally, Students Speaking Informally, and Quiet) over each curriculum area, it seems that teachers allow students to talk mainly during art, free time, and physical education. This may be one indication that teachers still feel that talk is not something one uses for serious learning but rather a time merely to chat. The one exception to this notion was the amount of Students Speaking Informally that occurred during language arts, written. The writing process as elaborated by Graves (1983) is a method for teaching writing that encourages students to work together and to serve as audiences for each other's writing. As many classrooms included in the study were engaged in this form of writing instruction, much of the language categorized as Stu-

dents Speaking Informally happened during the language arts, written period.

In examining the amount of total time over all talk structures spent in each curriculum area, it became apparent that language arts/reading consumed 38% of the week's activity time. Mathematics consumed 16% of the week's activity time. If transition times and housekeeping activities were combined, they consumed 14% of the week's activity time, while science consumed 4% of the time and social studies consumed 7%. It seems desirable that transition and housekeeping procedures become streamlined so that the extra time might be distributed over such curriculum areas as social studies and science. None of the teachers in the study had less than five years teaching experience, so it seems that this need to streamline peripheral activities is not symptomatic of only the beginning teacher.

If talk is so important for learning, why is not more of it going on in today's classroom? Is the old adage "A quiet classroom is a good classroom" still being given credence? Are teachers unaware of the talk structures that are happening within their classrooms? Do teachers view talk as only a way to socialize rather than a way to learn? Are teachers unaware of specific techniques for conducting brainstorming sessions and for structuring small group discussions? Whatever the cause, it seems important for researchers, teacher educators, and curriculum specialists within school districts to discover the answers to these questions so that practice is more aligned with recent theory and research.

Note

1. This study was funded by the Social Sciences and Humanities Research Council of Canada.

References

Allen, R., Brown, K., & Yatvin, N. (1986). *Language learning through communication.* Belmont, CA: Wadsworth.

Anderson, L., Evertson, C., & Brophy, J. (1979). An experimental study of effective teaching in first grade reading groups. *Elementary School Journal, 79,* 193-223.

Barnes, D. (1976). *From communication to curriculum.* Harmondsworth, UK: Penguin.

Barnes, D. (1990). Oral language and learning. In S. Hynds & D.L. Rubin (Eds.), *Perspectives on talk and learning* (pp. 41-57). Urbana, IL: National Council of Teachers of English.

Barnes, D., & Todd F. (1977). *Communicating and learning in small groups.* London: Routledge and Kegan Paul.

Beebe, W., Steven, T., & Masterson, J. (1982). *Communicating in small groups.* Glenview, IL: Scott Forsman.

Bellack, A., Kliebard, H., Hyman, R. & Smith, F. (1966). *Language of the classroom.* Columbia, MO: Teachers College Press.

Berliner, D., & Casanova, U. (1988). Are schools hampering language competency? *Instructor, 97*(7), 14-15.

Britton, J. (1970). *Language and learning.* London: Penguin Press.

Brophy, J., & Evertson, C. (1974). *Process-product correlations in the Texas teacher effectiveness study: Final report* (Research report no. 74-4). Austin, TX: Research and Development Center for Teacher Education, University of Texas at Austin. (ERIC Document Reproduction Service No. ED 099345)

Brophy, J., & Everston C. (1976). *Learning from teaching: A developmental perspective.* Boston, MA: Allyn and Bacon.

Brophy, J., & Good, T.L. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). New York: Macmillan.

Cazden, C. (1986). Classroom discourse. In M. Wittrock (Ed.), *Yearbook of research on teaching* (pp. 432-463). New York: Macmillan.
Chang, G.L., & Wells, G. (1988). The literate potential of collaborative talk. In M. Maclure, T. Phillips, & A. Wilkinson (Eds.), *Oracy matters* (pp. 95-110). Philadelphia, PA: Open University Press.
Chilcoat, G.W. (1987). Teacher talk: Keep it clear. *Academic Therapy, 22*, 263-271.
Chorny, M. (1981). A Canadian perspective: Focus on talk. *English Eduction, 13*(1), 32-35.
Cohen, E.G. (1989). Can classrooms learn? *Sociology of Education, 62*(2), 75-94.
Corson, D. (1984). The case for oral language in schooling. *Elementary School Journal, 83*, 265-285.
Cryan, J.R. (1984). Forms and functions of child-to-child interaction in classroom settings: A review. *Childhood Education, 60*(5), 354-360.
Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research, 13*(1), 9-19.
Edelsky, C., Draper, K., & Smith, K. (1983). Hookin' em in at the start of school year in a "whole language" classroom. *Anthropology and Education Quarterly, 14*(4), 257-281.
Edwards, A., & Furlong, V. (1978). *Language and teaching: Meaning in classroom interactions*. London: Heinemann.
Edwards, A.D., & Mercer, N. (1987). *Common knowledge: The development of understanding in the classroom*. New York: Methuen.
Edwards, A.D., & Westgate, D.P.G. (1987). *Investigating classroom talk*. London: Falmer.
Eeds, M., & Wells, D. (1989). Grand conversations: An exploration of meaning construction in literature study groups. *Research in Teaching English, 23*(1), 4-29.
Evertson, C., Anderson, L., & Brophy, J. (1978). *Texas junior high school study: Final report of process-outcome relationships* (Report No. 4061). Austin, TX: Research and Development Center for Teacher Education, University of Texas.
Evertson, C., Emmer, E., & Brophy, J. (1980). Predictors of effective teaching in junior high mathematics classrooms. *Journal for Research in Mathematics Education, 11*, 167-177.
Fisher, C., Marliave, R., Cahen, L., Dishaw, M., Moor, J., & Berliner, D. (1978). *Teaching behaviors, academic learning time and student achievement: Final report of Phase III-B, beginning teacher evaluation study*. San Francisco, CA: Far West Laboratory for Education Research.
Flanders, N. (1970). *Analyzing teaching behavior*. Reading, MA: Addison-Wesley.
Furst, N., & Amidon, E. (1967). *Interaction analysis: Theory, research and application*. Reading, MA: Addison-Wesley.
Gambell, T.J. (1980). *Talk contexts and speech styles: Planning for oral language*. Paper presented at annual meeting of Canadian Council Teachers of English, Halifax, NS. (ERIC Document Reproduction Service No. ED 193637)
Gilbert, J., & Pope, M. (1986). Small group discussions about conceptions in science: A case study. *Research in Science and Technological Education, 4*(1), 61-76.
Good, T., & Beckerman, T. (1978). A naturalistic study in sixth grade classrooms. *Elementary School Journal, 78*, 192-201.
Good, T., & Grouws, D. (1977). Teaching effects: A process-product study in a fourth grade mathematics classroom. *Journal of Teacher Education, 28*(3), 49-54.
Good, T., & Grouws, D. (1979). The Missouri mathematics effectiveness project: An experimental study in fourth grade classrooms. *Journal of Educational Psychology, 71*(3), 355-362.
Graves, D. (1983). *Writing: Teachers and children at work*. Exeter, NH: Heinemann.
Graybeal, S., & Stodolsky, S. (1985). Peer work groups in elementary schools. *American Journal of Education, 93*(3), 409-428.
Gunderson, L., & Shapiro, J. (1986). Some findings on whole language instruction. *Reading-Canada-Lecture, 5*, 22-26.
Hansen, J. (1987). Organizing student learning. Teachers teach what and how. In J. Squire (Ed.), *The dynamics of language learning* (pp. 321-335). Urbana, IL: NCTE.
Hennings, D.T. (1975). *Mastering classroom communication: What interaction analysis tells the teacher*. Pacific Palisades, CA: Goodyear.
Inner London Education Authority. (1984). *Improving secondary schools*. Report of the Committee on the Curriculum and Organization of Secondary Schools. London: ILEA.
Jones, M. (1988). *Lipservice: The story of talk in schools*. Philadelphia, PA: Open University Press.

King, A. (1989). Verbal interaction and problem solving within computer-assisted cooperative learning groups. *Journal of Educational Computing Research, 5*(1), 1-15.

Langer, S. (1960). *Philosophy in a new key* (3rd ed.). Cambridge, MA: Harvard University Press.

Lindfors, J. (1987). *Children's language and learning* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Logan, T. (1986). Cooperative learning: A view from the inside. *The Social Studies, 77*(3), 123-126.

McDonald, F., & Elias, P. (1976). *The effects of pupil performance on pupil learning. Final report. Beginning teacher evaluation study, Phase II, 1974-1976*. Princeton, NJ: Educational Testing Service.

Phillips, T. (1988). Why successful small-group talk depends upon not keeping to the point. In M. Maclure, T. Phillips, & A. Wilkinson (Eds.), *Oracy matters* (pp. 69-81). Philadelphia, PA: Open University Press.

Pigford, A.B. (1988). Teachers let students talk. *Academic Therapy, 24*(2), 193-197.

Platt, N. (1984). How one classroom gives access to meaning. *Theory Into Practice, 23,* 239-245.

Province of British Columbia Ministry of Education, Student Assessment Branch. (1988). *Oral communication in the primary grades: Teachers' resource package*. Victoria, BC: Author.

Province of British Columbia Ministry of Education, Program Development. (1989). *Primary program*. Victoria, BC: Author.

Rosenshine, B. (1976). Recent research on teaching behaviors and student achievement. *Journal of Teacher Education, 27*(1), 61-64.

Sharan, S., & Shachar, H. (1988). *Language and learning in the cooperative classroom: Recent research in psychology*. New York: Springer Verlag.

Slavin, R. (1983). *Cooperative learning*. New York: Longman.

Soar, R. (1977). An integration of findings from four studies of teacher effectiveness. In G. Borich (Ed.), *The appraisal of teaching: Concepts and processes*. Reading, MA: Addison-Wesley.

Soar, R., & Soar, R. (1972). An empirical analysis of selected follow-through programs: An example of a process approach to evaluation. *National Society for the Study of Education Yearbook, 71*(2), 229-259.

Solomon, D., & Kendall, A. (1979). *Children in classrooms: An investigation of person-environment interaction*. New York: Praeger.

Staab, C. (1990). Talk in whole language classrooms. In V. Froese (Ed.), *Whole-language: Practice and theory* (pp. 15-47). Scarborough, ON: Prentice-Hall.

Stallings, J., Cory, R., Fairweather, J., & Needels, M. (1977). *Early childhood education classroom evaluation*. Menlo Park, CA: SRI International.

Stallings, J., Cory, R., Fairweather, J., & Needels, M. (1978). *A study of basic skills taught in secondary schools*. Menlo Park, CA: SRI International.

Stallings, J., & Kaskowitz, D. (1974). *Follow-through classroom observation evaluation 1972-1973* (SRI project URU-7370). Stanford, CA: Stanford Research Institute.

Stallings, J., Needels, M., & Stayrook, N. (1979). *The teaching of basic reading in secondary schools, Phase II and Phase III*. Menlo Park, CA: SRI International.

Stewig, J.W. (1988). Oral language: A place in the curriculum. *Clearing House, 62*(4), 171-174.

Strickland, D., Dillon, R., Funkhouser, L., Glick, M., & Rogers, C. (1989). Research currents: Classroom dialogue during literature response groups. *Language Arts, 66*(2), 192-201.

Tikunoff, W., Berliner, D., & Rist, R. (1975). *An ethnographic study of the forty classrooms of the beginning teacher evaluation study known sample* (Technical Report No. 75-10-05). San Francisco, CA: Far West Laboratory for Educational Research and Development.

Trudge, J., & Caruso, D. (1988). Cooperative problem solving in the classroom: Enhancing young children's cognitive development. *Young Children, 44,* 46-52.

Vygotsky, L.S. (1962). *Thought and language*. Cambridge, MA: MIT.

Wasserman, S. (1989). Children working in groups? It doesn't work! *Childhood Education, 65*(4), 201-205.

Webb, N., Ender, P., & Lewis, S. (1986). Problem-solving strategies and group processes in small groups learning computer programming. *American Educational Research Journal, 23*(2), 243-261.

Webb, N., & Kenderski, C. (1984). Student interaction and learning in small-group and whole-class settings. In P. Peterson, L. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes* (pp. 153-170). New York: Academic Press.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 1, March 1991, 49-63*

FRANK H. ECHOLS, Jr.

CHARLES S. UNGERLEIDER

CAROL LA BAR

and

LEROI B. DANIELS
University of British Columbia

Students' Knowledge of the Charter of Rights and Freedoms and Their Agreement With Its Provisions

The notion of entrenching a charter of rights in Canada's Constitution was among the most hotly contested issues both prior to and after patriation. Those concerned with legal education and human rights could foresee that people's knowledge of their rights and their willingness to accord rights to other citizens might be affected by the inclusion of the Charter of Rights and Freedoms in the Constitution. This study was undertaken in order to determine students' knowledge of the Charter and the extent of their support for its provisions. A 27-item questionnaire was developed and administered to 3,161 students. These secondary students knew slightly more than 60% of the items. Knowledge and support were positively correlated ($r=.60$, $p<.001$). *Null findings are reported for parental social status, country of birth, first language, and ethnic identity. However, grade level, parental educational level, and source of information about the law were significantly related to knowledge and support. The results have educational implications, particularly for those interested in law-related education.*

In a free and democratic society, it is important that citizens know exactly what their rights and freedoms are, and where to turn for help and advice in the event that those freedoms are denied or rights infringed upon. In a country like Canada—vast and diverse, with 11 governments, two official languages and a variety of ethnic origins—the only way to provide equal

Frank Echols is an assistant professor of the sociology of education in the Department of Social and Educational Studies whose research interests include teacher education, school effects, and legal rights.

Charles Ungerleider, an associate professor of the sociology of education in the Department of Social and Educational Studies, is involved in research about intergroup relations and citizenship education.

Carol La Bar is a fellow of the Association for Values Education and Research. She has been involved in a variety of projects concerned with moral education, critical thinking, and law-related education at the University of British Columbia.

Leroi Daniels is a professor of the philosophy of education in the Department of Social and Education Studies, and a founding member of the Association for Values Education and Research. His research interests include moral education, law-related education, and global education.

> protection to everyone is to enshrine those basic rights and freedoms in the Constitution. Jean Chretien, Minister of Justice, 1982 (*Charter of Rights and Freedoms: A Guide for Canadians*, 1984, p. v)

In a democracy, a nation's constitution provides the basic framework for the legal system by defining the rights and obligations obtaining among the nation's people and between them and their government. Fundamental changes in a nation's constitution are typically accompanied by equally momentous changes in those relations. Thus, with the patriation of the Constitution in 1982, Canada embarked upon a new course in its legal and constitutional history.

The notion of entrenching a charter of rights in Canada's Constitution was among the most hotly contested issues both prior to and after patriation. Those concerned with legal education and human rights could foresee that people's knowledge of their rights and their willingness to accord rights to other citizens might be affected by the inclusion of the Charter of Rights and Freedoms in the Constitution. For that reason, the Law-Related Education Group at the University of British Columbia made application to the Social Sciences and Humanities Research Council of Canada to undertake a study that would provide foundational knowledge about people's knowledge of rights protected by the Charter of Rights and Freedoms and of their willingness to see these rights accorded to others.

Related Literature

Because the promulgation of the Charter of Rights and Freedoms is a somewhat recent event in Canadian history, the number of studies similar to the British Columbia Charter Survey is limited. We are aware of one other large scale survey that addresses this topic; however, differences in method limit comparisons of our respective findings. The Charter Project, undertaken by Sniderman, Fletcher, Russell, and Tetlock (1989), addresses the state of opinion in Canada, regarding the rights bestowed on its citizens by the Charter of Rights and Freedoms. Telephone interviews were conducted with a sample of 2,084 adults and a purposive sample of "political elites." The latter group has three components: 513 federal and provincial legislators, 260 officials in the federal Ministry of the Solicitor General, and a sample of 352 senior lawyers. In addition to the interviews, self-administered questionnaires were returned by 1,250 adults and 936 of the political elites. Their first report from this large data base focuses on the development of a "political theory of language rights" (Sniderman et al., 1989, p. 282).

Although they found "overwhelming support" for the principle of bilingualism by anglophones and francophones, the authors conclude that respondent attitudes toward language rights are shaped by a complex interplay of values and group identification. Their analysis identifies egalitarianism, rather than civil liberties, as a source of support for language rights. There is, nevertheless, a body of literature concerning the attitudes that Canadian students hold toward particular rights and freedoms. Kirkwood and Nediger (n.d.) surveyed Canadian students in grades 7 and 10 to determine their "level of general knowledge and attitudes toward Canada." Of the 100 items in their scale, six in their Special Issues Subscale were very similar to various

sections of the Charter. Their sample consisted of 10,821 students drawn from a random selection of a 1982 listing of public school boards representing all 10 provinces and the two territories.

Scores on the six items that are roughly equivalent to Charter provisions indicated that both grade 7 and 10 students were very supportive of the idea that Canadians should be able to live or work in any province; that Canadians should have the right to free speech; and that men and women should be paid equal wages for equal work. Seventy-one percent of grade seven pupils supported the right of Canadians to criticize government policies while 87% of those in grade 10 were supportive. Seventy-five percent and 73% respectively were supportive of French-speaking Canadians in all provinces having the right to education in French. Support was lower (69% and 64% respectively) to allow immigrants to keep their own culture and traditions.

English and French language rights have received rather more research attention than other Charter rights. The degree of support for bilingualism has been of substantial interest since before the adoption of the Official Languages Act in 1969. Johnstone (1969), in his 1965 national study of young people for the Royal Commission on Bilingualism and Biculturalism, found that Anglophones were "only moderately tolerant in their attitudes regarding bilingual goals for Canada." Predictably, the views of Francophone youth were more positive. Johnstone states, "Although they (Anglophones) were almost as ready as the Francophones to support bilingualism for Canada as an ideal, they were less likely to endorse positions which implied personal or social change"(p. xvi).

Bibby and Posterski (1985) conducted a national survey of 3,600 teenagers aged 15 to 19 years. Although much of the data from their 1984 Project Teen Canada survey are not relevant to present concerns, their data regarding bilingualism are relevant. They report that in every region of the country Canadian teens are more accepting of bilingualism than Canadian adults. Nationally, 71% of teens are in favor of the policy that Canada should have English and French as its two official languages, compared with 55% of adults surveyed in 1980-1981 and 49% surveyed in 1975. The lowest level of support among teens is in British Columbia (57%); among adults it is 44%, second lowest only to respondents from the prairie provinces.

Purposes of the Study

The above mentioned studies are a subset of a larger body of literature about the values that Canadians hold. And though they sometimes address issues pertinent to the Charter of Rights and Freedoms, they are generally focused on other issues and suffer from methodological shortcomings that limit their applicability and utility in addressing the questions that animated our investigation. The specific purpose of the British Columbia Charter Survey was to ascertain what British Columbia secondary students know about the legal rights included in the Charter of Rights and Freedoms and their attitudes toward those rights.

Method

This exploratory, descriptive study used a group administered questionnaire as the data collection instrument. The questionnaires were given to intact

classrooms by members of the research team. This procedure was followed to standardize the administration of the survey instruments.

Sample

A multistage cluster sample of 3,161 students was drawn by classifying all the communities in the province with secondary schools into six categories. Statistics Canada classifications for community size were used: 1,000-2,499; 2,500-4,999; 5,000-9,999; 10,000-29,999; 30,000-99,999; 100,000 and over. Two communities were then randomly selected from each category. In those communities with more than one secondary school, the schools were also randomly drawn. There were two exceptions to this practice. In those two situations we were required to use schools where we could gain cooperation from school administrators rather than use the two schools we had randomly chosen. All social studies classrooms from grade 8 to 11 were included in the survey. Due to a conflict between the timing of our data collection and provincial examinations, we excluded grade 12 classes from our sample. However, our sample does include 139 grade 12 students who were present when we administered the survey to what we had believed to be grade 11 classes.

Table 1 shows the relation between our sample of secondary school students and the population of secondary school students in British Columbia. Based on the relationship between the size of the student sample and the population from which it was drawn, the findings reported in this article should fall in plus or minus 1.6% 95 out of 100 times.

Instrumentation

The instrument developed for use in the survey was the result of extensive development and repeated pilot testing. Our pilot tests confirmed that the legal language of the Charter made it to difficult to phrase a question for each of the rights guaranteed by the Charter. We eliminated a number of items that proved to be misleading or confusing or both. The final version contained 27 vignettes representing selected rights granted by the Charter.

We do not have reliability data for the instrument, but we do have validity data. Two lawyers with expertise in constitutional law were asked to evaluate the instrument for content validity. They agreed that the instrument was a faithful representation of the content of the Charter of Rights and Freedoms. An instrument is considered to have concurrent validity if it differentiates between groups already considered different. We believe the questionnaire has concurrent validity. A sample of law students had significantly higher scores than a comparison group of university students. Both of these groups had significantly higher scores than a sample of students in grade 12.

Knowledge of and attitudes toward provisions of the Charter were ascertained using a format that involved the presentation of a vignette organized around one right or freedom. The following is illustrative of the presentation of the vignettes:

> A man went to the personnel department of a major bank and asked to apply for a particular position for which he was trained. He was told that he could

not apply for the position as he was not smart enough to take the next step and get a promotion.

After each vignette, respondents were asked a question to ascertain whether they knew the right in question. Thus, for this vignette, respondents were asked "Does the man have the legal right to be considered for the position?" and were given the opportunity to answer "yes," "no," or "don't know." To ascertain the respondents' attitude toward according the right to others, they were presented with an assertion designed to elicit their agreement or disagreement with the right. Thus, for the above vignette, respondents were presented with the assertion "The man should have the legal right to be considered for the position." This was followed by a five-point Likert scale with which they could indicate their agreement or disagreement with the assertion. This five-point scale was reduced to a three-point scale in the analysis. Responses were collapsed into "agree," "disagree," and "undecided." The students were asked to respond to 27 vignettes of this type presented in a paper and pencil format. The 27 vignettes are in the appendix.

Respondents were asked several background questions as part of the survey. Questions were asked to ascertain respondents' age, gender, grade level, first language spoken, language most frequently used at home, parental occupation, parental educational attainment, community size, and where they obtained most of their information about the law. Students also were asked their cultural background. The relevant question read: "Every person comes from a different kind of background in terms of family origin, culture, nationality and beliefs. In addition, each person may have their own sense of who they are. Please answer the following question to indicate what you feel is your own cultural background in terms of the cultural group to which you belong." Respondents were provided with examples such as English Canadian, Native Canadian, Japanese Canadian, and German Canadian and then were asked to indicate the cultural group to which they see themselves as belonging.

We wished to explore how these variables might explain variation in knowledge and attitudes. Such analyses might help us understand how underlying social processes are related to the outcome variables. If significant

Table 1
Grade Distribution of Sample and Population

	Population		Sample	
Grade Level	N	%	N	%
8	36,937	19.7	433	13.7
9	37,480	20.1	786	24.9
10	38,571	20.9	753	23.8
11	38,494	20.6	1,050	33.2
12	34,913	18.7	139	4.4
Totals	186,395	100.0	3,161	100.0

variations between social groupings were found, then the results could be of value in the development of target strategies to increase knowledge about the Charter.

In what follows we report how these background factors relate to how the students view the Charter of Rights. We also describe six clusters of rights and the degree to which they are known and supported.

Results

Table 2 contains a summary of the student responses organized into clusters of rights. As indicated in Table 2, students obtained a "correct" knowledge score of 61.9, which is the mean of the number of right answers calculated as a percentage of the total. They are most knowledgeable about their language rights and equality rights and least knowledgeable about their democratic and legal rights. The students' agreement with the rights roughly parallels their responses to the knowledge items. They are in greatest agreement with language rights and mobility rights and in least agreement with legal and democratic rights.

Gender

To determine if gender differences in patterns of socialization and education were likely to affect their knowledge and attitudes, respondents were asked to indicate whether they were male or female. Male secondary students are more likely to know the provisions of the Charter than are females. Their respective group means are 63.4 and 60.6. Although the males know the Charter better than females, the latter are more likely to agree with its provisions. Females have a mean agreement score of 62.4, which is significantly higher than the 61.4 agreement score for the males.

Age

Age has a significant impact on the extent to which students know the Charter and are in support of its provisions. Increases in age are associated with higher mean scores through age 18. Students 17 and 18 years of age have the highest scores (65.0), and the youngest age group, 12 through 14 years of age, have the lowest scores (56.1). Older students, 19 through 21 years of age, break the age-related trend with a decline to 59.8. This group is much smaller in size than the other age groups, with a group size of 64.

An increase in support for the Charter is also associated with increases in age. Students between the ages of 12 and 14 show the least agreement with the Charter as revealed by their mean score of 58.3. Each of the remaining age categories shows significantly higher levels of agreement with the Charter when contrasted with this youngest age group. Students who are 17 and 18 years of age indicate the greatest support for the Charter with respective scores of 63.8 and 63.7.

Grade Level

Grade level has a substantive impact on the levels of knowledge and agreement on the part of respondents as seen in Table 3. The grade 8 students are the least knowledgeable group and the grade 12 students are the most knowledgeable. Each succeeding grade has a mean knowledge score that is

Table 2
Percentage of Correct Responses and Percentage of Agreement with Charter Provisions by Clusters of Rights

	Mean % Correct	Rank	Mean % Agree	Rank
Mobility Rights	62.8	4	68.4	2
Language Rights	74.8	1	73.6	1
Democratic Rights	52.7	6	55.5	6
Fundamental Freedoms	66.9	3	60.2	4
Legal Rights	57.2	5	57.9	5
Equality Rights	67.5	2	64.5	3
Overall	61.9		62.0	
sd	15.3		13.2	

Item numbers seen in the appendix that correspond to each cluster of rights are as follows: democratic rights are items #7, #9, #15; equality rights are items #10, #11, #18, #23; fundamental freedoms are items #4, #8, #13, #17, #20; language rights are items #3 and #21 (Official Language Rights) and items #14 and #22 (Minority Language Educational Rights); legal rights are #1, #2, #5, #6, #16, #19, #25, #26, #27; and mobility rights are items #12 and #24.

significantly higher than that of the previous grade. The relationship between grade level and support for the Charter has a similar distribution to that of the knowledge scores.

Given the close association between age and grade, we wanted to know if grade had a separate effect from that of age. A significant ($F<.05$) interaction between grade and age does occur for both outcome measures. Inspection of the matrix of mean values for the two outcome variables revealed a consistent pattern; scores increased with grade level when age was held constant. The results confirmed that schools do have an educational impact on student knowledge of the Charter. Schools appear to have a liberalizing function as well.

Country of Birth, First Language, and Ethnicity

The Charter of Rights and Freedoms defines the rights and obligations among the nation's people and between them and their government. Knowledge of one's rights and duties empowers individuals, and such knowledge aids in the process of cultural integration for newcomers to Canada. Over the last 15 years, British Columbia has experienced an increase in the number of Asian, Sikh, and Latin American immigrants. We wanted to find out if individuals with different backgrounds varied in their knowledge of their rights and freedoms and in their agreement with the provisions of the Charter.

More than 90% of the respondents were born in Canada, with slightly fewer than 10% coming from other countries. Students who were Canadian-born have a lower knowledge score, 61.8, than those not born in Canada who have a score of 63.0. A similar difference of 1.1 percentage points separates

Table 3
Grade Level, Charter Knowledge, and Agreement with the Charter

Grade Level	n	% Correct	% Agree
Grade 8	421	56.7	58.5
Grade 9	770	58.9	59.5
Grade 10	764	62.8	62.9
Grade 11	1,043	64.6	63.5
Grade 12	145	68.3	68.4

the support scores for the two groups. Students born in Canada have a mean score of 61.8, which indicates a lower level of agreement with the Charter than the mean of 62.9 for students not born in Canada. None of the differences is statistically significant.

Respondents were asked to indicate the first language they had learned and the languages they most frequently used at home. Whether or not one speaks English as a native tongue is of little consequence in terms of the extent to which students know about or agree with the Charter. The total number of students who learned to speak a language other than English as a first language is slightly less than 10%. The mean knowledge score for anglophones is 62.0, which is almost identical to the 62.2 score for non-native speakers of English. The respective agreement scores for the two groups are 61.9 and 62.7. Again, the differences are not statistically significant.

A person's first or most frequently used language may not be an accurate indication of his or her cultural background. For this reason, we asked the students their socioethnic identification. Approximately 65% of the students identify themselves as being English-Canadian. This group has a mean score of 62.1, while those who identify themselves as members of other groups have a mean score of 62.6, a difference of 0.5. There is a similar difference between the two groups in terms of their extent of support for the Charter. Individuals who do not label themselves as English-Canadians report a mean score of 62.7, which is slightly higher than the score of 62.1 for English-Canadians. The differences between the means are not statistically significant.

Sources of Information About the Law

Students were asked to indicate where they obtained their information about the law. Those who cited television as their primary source are the least knowledgeable about the Charter, with a mean score of 59.5%. At the other end of the continuum are students who report school courses as their principal source of law information. Their average percent correct is 66.5. Differences among the mean scores between each of the categories were statistically significant.

Support for provisions in the Charter echoes the pattern revealed by the knowledge scores. Students who are least in agreement with the Charter's provisions report television as their primary source of information about the law and those most in favor of these provisions identify school courses as

their source. The respective means are 60.1 and 65.0. Each category mean is significantly higher than that reported for the television category.

Parental Educational Level

The data in Table 5 indicate that offspring of reasonably well educated parents have the highest scores on both outcome measures. With two exceptions, the magnitude of the difference between scores associated with mother and father is 0.3% or less. It is interesting to note there is little difference between the scores of students with parents who have some postsecondary education and those parents who have at least a university degree.

The results suggest that meaningful parental education distinctions with regard to student knowledge of the Charter occur at two points. These are: (a) between nongraduates of high school and graduates, and (b) between high school graduates and those with education beyond the secondary level. When one considers the column of figures under agreement with the Charter, one cutoff point emerges. That division takes place between high school graduates and those who are educated beyond secondary education.

Parental Social Status

Parental social status was measured by classifying each parent's occupation according to the socioeconomic scale developed by Blishen and McRoberts (1976). These scores were collapsed into six categories with Class I having the lowest ranking and Class VI having the highest ranking. Three fathers were listed as homemakers. Almost one third (909 or 32.2%) of the pupils who answered this question listed their mothers as homemakers.

Respondents who listed their mothers as homemakers report a knowledge score of 62.0 and an agreement score of 61.5. These scores are higher than the corresponding scores for the lowest social status category, which are 61.0 and 61.4, but are below the values of the other five categories. Their knowledge values range from a low of 62.6 for Class II to a high of 63.3 for Class IV and Class V. The agreement with the Charter scores range from a low of 62.5 for Class II to a high of 63.7 for Class IV and Class VI. Although the highest group means are significantly different from the lowest group means, there is no main effect for mother's social status on either dependent variable.

Table 4
Law Information Sources, Charter Knowledge, and Agreement with the Charter

Source of Information	n	% Correct	% Agree
Television	818	59.5	60.1
Friends/Family	1,351	61.6	62.0
Newspapers/Magazines	389	64.4	63.4
School Courses	417	66.5	65.0

Table 5
Parental Educational Level, Charter Knowledge, and Agreement with the Charter

Level of Education[1]	n	% Correct	% Agree
Some High School			
Mother	820	61.0	61.1
Father	1,035	60.8	61.4
High School Graduate			
Mother	982	61.7	61.4
Father	811	62.0	61.5
Some Post-Secondary			
Mother	749	63.5	63.6
Father	674	63.6	62.8
University Degree or more			
Mother	405	63.6	63.7
Father	444	63.0	63.6

[1]The educational level for each parent was statistically significant for each outcome variable ($F < .01$).

Father's social status follows the same pattern: no main effect, but statistical significance is obtained between the highest and the lowest group means. Students whose fathers' occupations are ranked in the top category, Class VI, have the highest knowledge score (63.7). This score is significantly different from Class II, which is 61.3. Classes I, III, and IV have almost identical scores of 61.8, 61.7, and 61.9 respectively. Similar results are found with the agreement scores. Pupils with a father whose occupation places them in Class VI have an agreement score of 64.0. This value is significantly higher than the scores of 61.4 that are associated with students whose father's occupation is ranked in Class I or Class II.

Other Relationships

The correlation between knowledge and attitudes was calculated. The positive correlation ($r = .595$ $p \leq .001$) was strong and significant, indicating that knowledge of rights and willingness to accord rights to others are associated.

Summary Profiles

A summary of the findings is presented as a series of profiles. These profiles are based on comparisons between the overall sample and the distributions of individuals whose scores are approximately one standard deviation above or below the mean of the two dependent variables. These are general descriptions and are not meant to be considered as exclusive categories. There are exceptions to each of the profiles. There are four profiles: the individual student most likely to know the provisions contained in the Charter, the student least likely to know those provisions, the individual most likely to

agree or support the Charter provisions, and the individual least likely to agree or support the rights enshrined in the Charter of Rights and Freedoms.

The student who is most likely to know the provisions of the Charter is male, aged 17 or 18, presently in grade 10, 11, or 12, and obtains most of his information about the law from school courses. The parents of this individual have attained an educational level beyond that of high school graduate. This individual is likely to have a mean knowledge score of 75% or above.

The student least likely to know the provisions of the Charter is female, presently in grade 8 or 9, and reports that her primary source of information about the law is television. The parents of this individual have either some secondary school education or have graduated from secondary school. This individual is likely to have a knowledge score of 50% or below.

The individual most likely to be supportive of the provisions in the Charter is 16, 17, or 18 years of age, presently in grades 10 through 12, whose primary source of information about the law is school courses. This student may be either male or female. His or her parents have continued their education beyond that of secondary school graduate. This individual is likely to have an agreement score of 75% or above.

At the other extreme of the support continuum we have those respondents who are least supportive of the Charter of Rights and Freedoms. The typical respondent is male, aged 12 through 14, presently in grades 8 or 9, and states that television is his primary source of information about the law. The parents of this individual either graduated from secondary school or obtained some education at the secondary school level. This individual is likely to have an agreement score of 50% or below.

People acquire knowledge of their rights and develop attitudes toward those rights and toward the extension of those rights to others simply as a consequence of the process of growing up in a particular society. It is nevertheless the case that the patterns of socialization differ among individuals and between groups. We have examined our data in an attempt to understand how people's knowledge and attitudes are affected by their differential exposure to different sources of information about rights.

Our results are compatible with a belief in Canadian society as an egalitarian society. Our data indicate that country of birth, first language, ethnicity, and parental social status are not significantly related to either knowledge of the Charter or agreement with its provisions. Although we risk going beyond our data, the significant and positive impact of education on both dependent variables suggests that formal education plays an instrumental role in these null findings. Parental educational attainment, student educational attainment, and school courses as a source of law information explained significant variations in both dependent variables.

The existence of protected rights and freedoms is insufficient to their guarantee. The exercise of one's rights and freedoms is dependent on knowledge of them. Our student sample "knows" slightly more than 60% of the items we have used to measure Charter knowledge. Judgments about the adequacy of this knowledge are arbitrary. We have no basis for comparison, although our study does provide a baseline for future contrasts.

In light of the Meech Lake Accord, it is worth noting that our sample of students knew the items associated with language rights significantly better than any other cluster of rights. Their support for these rights paralleled their level of knowledge.

Comparisons between clusters of rights is mitigated by the number of items in each cluster and their relative degree of difficulty. That said, the mean scores for legal and democratic rights are significantly lower than the corresponding scores for the other clusters. This information should be of value to those concerned with planning educational experiences and design of curriculum materials.

Ignorance of one's rights increases the likelihood that those rights will be infringed on. A person ignorant of rights is also more likely to infringe on the rights of others. Knowledge, as we have already shown, is associated with willingness to accord rights to others. Legal systems cannot function well without consensus and support. Respect for the rule of law and the rights of others is necessary to the administration of justice, and knowledge is a prerequisite to respect.

Implications for Education

Full citizenship implies knowledge and consideration of the legal principles that ought to guide action, an understanding of one's intentions, and the consequences that would result from pursuing a particular course of action. Study of a nation's constitution is fundamental to the assumption of full citizenship. In other words, citizenship involves the obligation to be aware of and observe one's legal obligations. One consequence of failing to educate people about their rights and freedoms is a reduction in their autonomy. People cannot take a full and active part in the affairs of their communities if they lack knowledge of the duties and obligations of citizenship.

We believe that education should include the development of understanding of the basic rights to which all people are entitled. Such understanding would necessarily include at least learning that there are different kinds of rights, that rights protect people by establishing what other people must, may, or may not do, and that rights are essential to an open society. One way of expressing the notion is that all rights impose duties on people to do certain things or the obligation not to interfere with the actions of others. Students may understand the impact that rights have by considering such duties and obligations. The Charter offers a marvelous opportunity for such education.

References

Bibby, R.W., & Posterski, D.C. (1985). *The emerging generation: An inside look at Canada's teenagers*. Toronto, ON: Irwin.

Blishen, B.R., & McRoberts, H.H. (1976). A revised socioeconomic index of occupations in Canada. *Canadian Review of Sociology and Anthropology, 13*, 71-79.

Charter of rights and freedoms: A guide for Canadians. (1984). Ottawa: Minister of Supply and Services.

Johnstone, J.C. (1969). Young people's images of Canadian society. *Studies of the Royal Commission on Bilingualism and Biculturalism*. Ottawa: Queen's Printer.

Kirkwood, K.J., & Nediger, W.G. (n.d.). *A survey of elementary and secondary pupils: Their knowledge and attitudes regarding Canada*. London, ON: Faculty of Education, University of Western Ontario.

Sniderman, P.M., Fletcher, J.F., Russell, P.H., & Tetlock, P.E. (1989) Political culture and the problem of double standards. *Canadian Journal of Political Science, 22,* 259-284.

Appendix
Knowledge Items: Charter of Rights Survey

1. In a small town a man is charged with three armed robberies. He demands a trial by jury. The community argues that it would be a waste of time and taxpayers' money to hold a jury trial.

 Does the man have a legal right to a trial by jury? (Yes)

2. The police arrived at Mr. J's apartment with a warrant to search for stolen cameras. After a brief search, they found a number of cameras in a kitchen cupboard. He was charged with possession of stolen goods and taken to the police station. At the station he was strip searched.

 Do the police have a right to strip search Mr. J? (No)

3. André and Marie live in Trois Rivieres, Quebec. They are driving across Canada and are now in Manitoba. They need information about the weather, but do not understand English. They call Environment Canada in Winnipeg, expecting to receive information in French.

 Do André and Marie have a legal right to receive information in French? (Yes)

4. A group of people wants very much to stop Canadian companies from building parts which will be used in nuclear missiles. Every year they demonstrate in the public park beside a factory which supplies parts for nuclear missiles. This year, the factory manager insisted that the group move away from the park.

 Does the factory manager have a legal right to prevent the group from demonstrating in the park? (No)

5. A man was charged with assault. He was found not guilty at his trial. A short time later a witness came forward with additional evidence suggesting he was guilty.

 Does the state have a legal right to bring the man to trial again? (No)

6. A police officer stopped a car driven by a woman and made a demand for a breath sample. The woman accompanied the police officer to the station but refused to take the test until she had the opportunity to call a lawyer.

 Does the woman have a legal right to refuse to give the breath sample before having an opportunity to call a lawyer? (Yes)

7. The elected members of the governing party of a provincial legislature (parliament) all agreed to postpone the sitting of the legislature for two years. The opposition party members disagreed with the action. They said that the governing party could not postpone the legislature for that long.

 Do elected governing party members have the legal right to postpone the sitting of the legislature for two years? (No)

8. A group of people wants to publish a newspaper. The group want to publish material arguing that people should not be allowed to vote unless they have high intelligence.

 Does the group have a legal right to publish such a newspaper? (Yes)

9. Mr. K. recently received his Canadian citizenship. The day after he received his Canadian citizenship it was announced that there would be an election in thirty days. Mr. K. decided he wanted to run for a seat in Parliament as an independent citizen.

Does Mr. K. have a legal right to run for Parliament? (Yes)

10. A religious group wishes to build a temple in a small town. They are willing to locate it at any number of places and there are no by-laws which prevent it. Many of the townspeople do not wish to have a temple belonging to this particular group in their town.

 Do the townspeople have a legal right to stop the religious group from building a temple? (No)

11. The RCMP established a policy to increase the number of Native Indians in the force. One-quarter of the trainees they accepted were to be natives. Several Caucasians (white people) who met the qualifications did not get in because of this policy. They claimed to have been discriminated against.

 Does the RCMP have a legal right to institute this policy? (Yes)

12. A Hungarian woman immigrated to Canada in 1956 and obtained Canadian citizenship in 1981. She went back to Hungary in 1983 and stayed for fifteen months.

 Does the woman have a legal right to come back to Canada under the same conditions as someone born in Canada? (Yes)

13. A teacher in a small town wanted to join a political party. The principal said she could not join the party as teachers were supposed to be neutral about politics.

 Does the principal have a legal right to stop teachers from joining a political party? (No)

14. Jean and Yvette Picard are Canadian citizens. They lived in Laval, Quebec and spoke only French until they moved to Vancouver. Their daughter, Giselle, was born in B.C. Jean and Yvette wish to send Giselle to a school where classes are taught in French.

 Do the Picards have a legal right to send Giselle to a public school where classes are taught in French? (Yes)

15. Ms. W. has lived in Canada for twenty years without taking out Canadian citizenship. She has decided to run as a member of Parliament in her area.

 Does Ms. W. have a legal right to run as a member of Parliament? (No)

16. Mr. S. was caught by two policemen while he was breaking into a hardware store. He was charged with breaking and entering. Mr. S. is deaf and wishes to have an interpreter at his trial.

 Does Mr. S. have a legal right to an interpreter at his trial? (Yes)

17. A small group of parents wish to protest a school policy. They intend to hold a demonstration outside the school. People living near the school want to stop the group from demonstrating.

 Do the parents have a legal right to hold the demonstration? (Yes)

18. A man went to the personnel office of a major bank and asked to apply for a particular position for which he was trained. He was told that he could not apply for the position because he was not smart enough to take the next step and get a promotion.

 Does the man have a legal right to be considered for the position? (Yes)

19. A woman was charged with speeding on the Trans Canada highway. She said that she was not guilty and wanted to be tried by a jury.

 Does the woman have a legal right to a trial by jury? (No)

20. An elementary school teacher decided that her class should say the Lord's Prayer every morning. The parents of two students refused to let their children participate. They said they wanted their children excused from the room when the Lord's Prayer was being said.

 Do the parents have a legal right to have their children excused from the classroom when the Lord's Prayer is being said? (Yes)

21. Jeanne Belanger is from New Brunswick. She wants to find out about employment opportunities in Vancouver. Jeanne goes to a Canada Employment Centre. She speaks French and cannot understand or speak English. She expects to have her questions answered in French.

 Does Jeanne Belanger have a legal right to have her questions answered in French? (Yes)

22. A school board introduced a new policy. The policy said that only children whose parents were educated in French in the school district may attend the French schools in the district. All other children must attend the English schools in the district.

 Does the board have the legal right to establish this policy? (No)

23. A female kitchen helper was working in a restaurant in the Parliament buildings. Her boss expected her to do work that he did not expect the male workers to do. Male and female workers were paid the same and had the same job classification.

 Does the female worker have the legal right to be treated the same as the male workers? (Yes)

24. Mr. B. has just moved to British Columbia from Ontario. He has applied for a job as a customs officer. A group of unemployed people march on government offices demanding that people from other provinces not be allowed to apply for jobs in B.C.

 Does Mr. B. have a legal right to be considered for employment on an equal basis with British Columbians. (Yes)

25. Mr. M. was a guard at a provincial jail. He was charged with hit and run driving. He said he didn't do it, but the warden fired him.

 Does the warden have a legal right to fire Mr. M? (No)

26. A police officer came to a young woman's apartment with an order for her to appear in court. The woman had 23 parking tickets which she had not paid. The woman answered the door and the officer handed her the order. Then he said he would like to search her apartment. The woman protested.

 Does the woman have the legal right to refuse to allow the police to search her apartment? (Yes)

27. An RCMP officer drove up to a bus stop and beckoned to a man standing there. The man went over to the car. The officer asked him his name and address. After answering these questions, the man asked the officer why he wanted this information. The officer did not reply and asked the man to get into the car. The man said he would not get into the car unless the officer told him the reason.

 Does the man have the legal right to refuse to get into the car? (Yes)

The Alberta Journal of Educational Research *Vol. XXXVII, No. 1, March 1991, 65-86*

NELLY McEWEN
Alberta Education

and

HAU CHOW
Alberta Hospital Ponoka

Issues in Implementing Indicator Systems

An indicator system is a tool for reform and improvement. Existing accountability systems often present indicators in isolation. Negative impacts can be minimized if an indicator system is perceived as an integrated whole. Two examples of a new generation of promising integrated indicator systems are identified. Implementation issues can be categorized into educational and social, administrative and technical, and political and financial. Resolution of issues often involves compromise among competing perceptions and values. Implementation strategies that incorporate multiple goals, indicators, methods, levels of analysis, and participants are proposed.

This article presents a definition and description of educational indicator systems and outlines their uses, guiding principles, and criteria. It identifies two developmental projects that are examples of a new generation of indicator systems that promise to provide an explanatory framework for education. They are particularly relevant to the Canadian context because they explore local, provincial, national, and international perspectives on education. The paper subsequently addresses the major issues related to the implementation of indicator systems, based largely on the difficulties that have manifested themselves in the American context where indicators have been corrupted in attempts to produce short-term gains to satisfy annual accountability demands in an essentially long-term endeavor. It classifies the issues into three categories: educational and social, administrative and technical, and political and financial, and discusses their implications. Positive and negative effects are identified and strategies proposed for minimizing the negative impacts. The final section proposes strategies for resolving the issues.

Indicator Systems

The current educational reform movement has led to greater accountability and an increase in monitoring and evaluation of schools and systems. A critical component is information. In order to determine what type of information to collect, many organizations are developing and implementing

Nelly McEwen is Assistant Director, Educational Standards and Indicators, in the Corporate and Fiscal Planning Branch. She is interested in research methodology and evaluation.

Hau Chow, formerly with Alberta Education, is now a psychologist at Alberta Hospital Ponboka. Her interests are in the areas of assessment and research methodology.

indicator systems. Because education is a complex enterprise that cannot be described by means of any single component, an educational indicator system should address this complexity by simultaneously examining interrelated aspects of the enterprise rather than a number of discrete parts. Such a system should provide measures of various components as well as information about how they work together to produce changes in the condition of the system over time (Odden, 1990). The context, input, process, product (CIPP) model holds considerable promise for an indicator system. Stufflebeam and Webster (1988) described the objectives, methods, and relation to decision making in the change process of each of the four types of evaluation based on the CIPP model. Because educational institutions exist fundamentally to foster human growth and development, assessments of students' needs must take into account relevant developmental variables, and assessments of institutions must take into account the components and activities that bear on meeting the needs of their students.

Definition and Description

An indicator system is a tool for reform and improvement. Accountability is one aspect of this endeavor. The information provided by an indicator system must be meaningful for diverse groups because education is a shared responsibility among the provincial government, the family, school jurisdictions, schools, the academic community, and other institutions. Those making decisions about educational matters and those responsible for public funds must be held accountable for their decisions. Indicators can paint a broad picture of the conditions of education and stimulate thinking about potentially effective policies (Shavelson, 1988, p. 6). Although indicators cannot describe a system completely, they can inform planning, policy, and decision making (McEwen & Zatko, 1989, p. 7). The goals, or intended benefits, of implementing indicator systems are to improve education, to provide a mechanism for accountability, and to assess the efficiency and effectiveness of the educational enterprise. An educational indicator system should provide:

- information that is feasible to gather, valid, and useful for policy decisions;
- logical clusters of indicators to increase the system's interpretative power;
- points of reference such as a previous value, the value of a comparison group, or some socially determined standard;
- measures of the ubiquitous features of schooling that are enduring, easily understood, feasibly measured, and generally accepted as valid and reliable statistics; and
- a reporting system that is accurate and timely (Oakes, 1986; Selden, 1988).

An indicator system is not synonymous with a student testing or assessment program. Much of the criticism leveled against the existing American accountability systems is due to their use as programs of student assessment without regard for the consequences or effects that their introduction produces. An indicator system should be a much broader, more comprehensive undertaking that attempts to provide a more balanced picture of the operation and outcomes of schooling, particularly if the indicators focus on a

broader range of desired outcomes. By focusing on student learning, and interpreting the findings in terms of educational conditions, better decisions may result from the enhanced information (McEwen, 1990, p. 13). Student outcomes are but one, albeit vital, component in this larger picture of educational conditions. To examine student outcomes in isolation, without taking into account the context and operation (inputs and processes) of schooling, denigrates a large body of research on the relationships among achievement and other variables, and may perpetuate poor pedagogical policies and practices.

The existing American accountability indicator systems were precipitated by the release of *A Nation at Risk* and other reports in 1983 that identified a dearth of information about schools and schooling (Selden, 1988). According to Finn (1988), the reform movement in the United States arose from widespread anxiety that deteriorating educational performance was endangering the spirit, economic vitality, technological prowess, and perhaps even the national security of the United States, while eroding the ability of individuals to have a happy, rewarding, safe, and virtuous life (p. 116). The reform movement insisted on tangible evidence of results, most commonly in the form of rising scores on various tests and other indicators of cognitive learning. This resulted in a high degree of homogenization or standardization of education, with a commensurate narrowing of the range of choice given to students and the breadth of discretion given to teachers. According to the Office of Educational Research and Improvement (OERI, 1988), 46 states have accountability systems of which 35 are state systems, two are local, and nine are a mixture of both state and local systems (pp. 28-29). The major types of indicators reported are student achievement on standardized testing, attendance, and completion of school. While 45 states collect background data, only 21 reported using these data for interpreting performance indicators. Large-scale testing programs form the core of the state accountability systems and often drive the local response. Questions about the current status and use of standardized tests are being raised increasingly by researchers, the media, and the public. The most common criticisms include the inability of existing tests to measure the full range of achievement; the limited overlap between the substance of standardized tests and the skills and content stressed in textbooks; a general narrowing of the curriculum as a result of the tests; emphasis on basic skills; and the confusing messages transmitted to policy makers and the public as a result of the long intervals between revising norms (Brandt, 1989; Koretz, 1989; Meisels, 1989; OERI, 1988).

Promising Projects

A new generation of indicator systems is currently under development. These new systems have the potential to reduce the negative impact of the American accountability systems that focus almost exclusively on student outcomes and have resulted in negative consequences for students and educational institutions. Two examples illustrate a new and broader perspective on indicator systems that attempt to embed student outcomes in the context of the educational system provided through schooling. Their major focus is to improve education.

The Organization for Economic Cooperation and Development (OECD) has completed its 18-month exploratory phase and is currently in the developmental phase of the International Educational Indicators (INES) Project that will conclude December 1991 (OECD, 1989a). Originally INES was divided into five networks, each of which was the responsibility of a particular country. These were:

1. enrollment, career paths and school leavers (Australia);
2. education outcomes (United States);
3. the functioning of schools (France);
4. assessing costs and resources (Austria);
5. attitudes and expectations (Netherlands).

Canada participated in networks 1 and 2 that included 20 and 15 countries respectively. The next phase of INES will finalize the conceptual and analytical framework for the production of indicators and delineate an indicator system. This international project adopted a systematic scheme as a framework for the development and interpretation of indicator systems: environment, resources, processes, and effects (Nuttall, 1990). An understanding of the effects or outcomes of education must be informed by the educational processes employed and resources (fiscal and human) deployed, against the background of contextual factors in the environment of schools or of education systems. The INES project is important because of its international scope and broad intents. It is exploring the major conditions of education that need to be examined in concert with desired outcomes if true educational improvement is to occur.

The second example is Alberta's Educational Quality Indicators (EQI) initiative. This is a collaborative endeavor between Alberta Education and 12 school jurisdictions to develop indicator systems to measure the success of the educational enterprise in the province. This three-year initiative has three components:

1. indicator systems (meeting three criteria: an interpretative framework to describe variation among students and schools; student outcomes related to the educational enterprise; and points of reference for comparing results);
2. methods (to collect, analyze, and interpret the information); and
3. reporting and dissemination (to inform diverse audiences of the results).

A four-dimensional model of education was developed to guide this initiative. It consists of partners (schooling, family, society), conditions (context, inputs, processes), student outcomes (cognitive, affective, behavioral) and time (grades 3, 6, 9, 12) (McEwen & Zatko, 1989). Each participating school jurisdiction is currently developing and will implement a local indicator system that represents its educational goals and priorities and reflects the local community's expectations. The information generated from the 10 concurrent collaborative action research projects will assist Alberta Education to develop provincial indicators. The interpretation and recommended directions of the local indicator projects, together with other provincial initiatives, is expected to provide a solid foundation for the implementation of an efficient and effective information system that measures the success of the educational enterprise in the province (McEwen, 1990).

The OECD and EQI projects represent a new generation of indicator systems because they are based on models of education that attempt to describe the interrelationships among components and permit multiple comparisons. Both recommend multiple data sources and multiple methods of analysis. Indicator systems, whether local, provincial, national or international, are interdependent because they reflect the goals and priorities of their originators. Education is a social mechanism for promoting shared values and societal goals. The common goal is an educated citizenry to deal with the increasing complexity of life in an interdependent world. A shared priority is the development of the next generation as efficiently and effectively as possible given competing demands for limited resources. It is both a social and economic imperative that students be prepared to contribute to their society (community, province, country) because we all live in the global village.

Issues

The implementation of any indicator system raises many issues about education. Even an exemplary indicator system is not a panacea for improving schooling. Education is influenced by interacting political, social, economic, and cultural forces. Kanter (1989) discussed how the turbulent, fast-paced environment of global competition and constant innovation in technology and corporate structure has dramatically changed our lives. Public, private, and independent institutions are initiating a wide variety of collaborative, multiorganizational approaches to address complex issues and problems. Issues management alliances (Austrom & Lad, 1989), known as public-private partnerships, innovating organizations, community development corporations, and self-regulatory agencies are an attempt to respond to the turbulent sociopolitical environment. Issues management is the proactive process of identifying and shaping social issues and emergent trends likely to affect an organization, as well as deployment of the necessary resources toward their effective and timely resolution (p. 23). According to Luke (1986), interconnectedness and interdependence have emerged as the most salient characteristics of the contextual environment of contemporary society. In order to deal with issues, macrolevel analysis, a multisector perspective, longitudinal research, and multidisciplinary approaches are required (Austrom & Lad, 1989, p. 36).

Schooling is a powerful institution for developing the next generation, but its support depends on the perceived value of its utility and effectiveness. The next sections present the issues for each of the three major areas: educational and social, administrative and technical, and political and financial and discuss their implications. Both positive and negative effects are identified. In many cases issues result from opposing ideas of educational excellence, equity, effectiveness, and efficiency.

Educational and Social Issues

Many issues that reappear are value conflicts. When economic, social, and demographic changes create social turmoil, public opinion shifts (Cuban, 1990, p. 8). Differences, as they become transformed by the media and political coalitions into pressure on schools to change, are dilemmas that

require political negotiation and compromises among policy makers and interest groups. Cuban (1990) stated that people turn to schools in times of social turmoil because dominant social groups want public schools to work on national ills, rather than risking major dislocations in society, by directly addressing major social problems. People have a shared, enduring belief that schools promote social mobility, creating national harmony and building solid citizens.

Context and Values

Education does not exist in a vacuum. It is important to understand the context in which schooling takes place. Changes in the economic, social, and cultural aspects of society have an impact on children's social, emotional, and cognitive development. The issue becomes how schooling can accommodate differences in children's learning potential as a result of the constantly changing society. Cuban (1990) conceptualized the struggle over reform as a conflict over values. When conservative values were prevalent (in the '20s, '50s, and '80s), schools were concerned with producing individuals who could compete: hence high academic standards, orderliness, efficiency, and productivity were prized in schools. When liberal values dominated (in the early 1900s, '30s, and '60s), concerns for minorities, the poor, and outsiders prompted school reforms to broaden student access to programs, linked schools to work in the community, and reduced academic achievement gaps between groups of students. Each political turn of the cycle left a residue in schools when the rhythm shifted. Indicators need to describe these trends in the context because they are important for interpreting results.

Canadian society has changed enormously over the past 25 years (Health and Welfare Canada, 1989). These changes have placed tremendous pressure on families as they try to prepare their children for an increasingly complex and fast-paced society. The traditional two-parent family with the mother staying at home has been largely replaced by dual-income families that have steadily increased since the 1960s from 25% to over 60%. The number of one-parent families has increased significantly over the past 25 years. The ethnocultural structure of Canada is also changing due to shifts in immigration patterns. Immigration is far more diverse now than it was a quarter of a century ago in terms of the national, linguistic, religious, and racial backgrounds of immigrants (Health and Welfare Canada, 1989, p. 32). The shifting patterns have a direct and immediate effect on the educational system. The growing number of non-English/non-French-speaking school-aged children places pressures on the educational system for increased resources (staff, time, and money) for language training. The increasing multicultural aspect of Canadian society also demands that the educational system address these complexities.

Both home and school environments are important in the conceptualization of student learning. Walberg (1984) identified home environment as one of the causal factors in student learning. In his synthesis of over 3,000 studies of student learning, Walberg concluded that the "alterable curriculum of the home" is twice as predictive of academic learning as family socioeconomic status. This "curriculum" refers to "informed parent-child conversations about school and everyday events; encouragement and discussion of leisure

reading ... expressions of affection and interest in the child's academic and other progress as a person" (p. 25). Fraser, Walberg, Welch, and Hattie (1987), in identifying the salient facets of student learning, stated that "the pupil is part of a home, school, and classroom and is subject to the influences from peers, teachers, parents, and the media" (p. 193). In other words, the amount of learning any individual child can demonstrate is partially determined by contextual factors outside the control of the educational system. Although a child's previous achievement level remains one of the most important variables in determining student outcome, that level is influenced by the child's home environment.

In order to attribute schooling effects to learning, techniques to control for background influences are required. Nagy, Drost, and Banfield (1985), in an investigation of community isolation and its influence on educational achievement, used path analysis to control for economic and parental effects. They found a substantial contribution for some of the isolation variables in accounting for achievement differences. Although background variables are not readily amenable to manipulation by educators, they have a great impact on achievement effects and need to be examined.

Curriculum and Assessment

In their documentation of the Canadian component of the Second International Science Study, Connelly, Crocker and Kass (1985) described the diversity in education in Canada in policy, organization, styles of teaching, curricular emphases and development, language of instruction, and in rights to instructional services (p. 311). This diversity introduces problems when the implementation of national indicators is contemplated.

Although school jurisdictions, provincial governments, and the public place much emphasis (testing is usually funded) and credibility (schools are held accountable) on test scores, this is not translated into instructional practice because teachers see this type of testing as an unnecessary intrusion into the classroom. Probably the most visible and most common indicator of student outcome or performance has been some type of test score. Assessment and its impact on the curriculum, the learning process, and the educational system as a whole has been the focus of several recent special issues: *Educational Leadership* (April 1989); *Phi Delta Kappan* (May 1989); *Educational Researcher* (December 1989); and the *Alberta Journal of Educational Research* (March 1990). The impact of the use of standardized tests on the curriculum is related to the use of these test scores as accountability tools (e.g., Koretz, 1989; Neill & Medina, 1989; Wiggins, 1989). The pressure to raise test scores has resulted in a disruption of the learning process because the priority becomes raising short-term test scores for the class rather than ensuring that students learn. This can result in practices such as teaching to the test and a general narrowing of the curriculum, where the curricular content is dictated by the content of the test (Shepard, 1989). Although test developers attempt to use current curricula to guide test construction, the resultant test may not be responsive to changes in curricula over time and differential emphasis across different systems. McLean (1985), in his survey of student evaluation in Canada, found that district and provincial achievement testing is common, but that few uses of the results were found; teachers

evaluate students, but do not communicate results widely; employers are more concerned with attitudes and behavior than marks; and postsecondary institutions are concerned about the comparability of school marks in the absence of common examinations (p. 47). In questioning the need to examine outcomes, Nagy (1990) concluded that testing technology distorts the curriculum; political realities interfere with doing the job properly; and people want to see "hard" numbers rather than anecdotal reports, even though these hide the real data and threaten people. Moreover, evidence suggests that the intended, translated, and achieved curriculum are different (e.g., Connelly, Crocker, & Kass, 1989).

Hargreaves (1989) recommended the following strategies for reforming curriculum and assessment: decentralization of curriculum development; administrative support for a collaborative teacher culture; mandatory guidelines requiring a broad and balanced curriculum and reinforced through the power of inspection; and a revamped assessment system designed to provide teachers with improved feedback about their students and their progress as a basis for curricular renewal (p. 170).

Structure of Schooling

Schooling consists of many conditions that are alterable. Indicator systems need to incorporate these conditions because information derived from such input and process indicators can suggest possible intervention strategies for improving education. Tyler (1984) said that "education ... is a social enterprise seeking to help persons acquire understanding, skills, attitudes, interests and appreciation" (p. 29). The structure of schooling needs to be examined if educational improvement is to occur. Some of the models of education that contribute to our understanding of the relationships between student learning and development and the conditions of schooling follow.

Carroll (1963) described the degree of learning as a function of time spent (determined by opportunity to learn and perseverance) over time needed (determined by aptitude, ability to understand instruction, and the quality of instruction). Hymel (1988) expanded Carroll's model to include four additional theories: Bloom's mastery learning model; Edmonds'/Brookover's effective schools correlates; Spady's premises of outcome-based schooling; and Hunter's mastery teaching concepts. Since 1975 Walberg and his colleagues have been trying to develop a comprehensive framework for the analysis of productivity and to test it in a variety of classroom studies (Walberg, 1984, pp. 20-21). Nine factors require optimization to increase student (affective, behavioral, cognitive) learning: student aptitude (ability, development, motivation), instruction (quantity and quality), and environment (home, classroom, peer group, mass media). Oakes' (1986) comprehensive model of the educational system consists of inputs, processes, and outputs. Shavelson, McDonnell, Oakes, Carey, and Picus (1987) developed a framework for a possible indicator system that piggybacks onto ongoing data collection through the American National Assessment of Educational Progress. Their piggyback indicator system looks at inputs, processes, and student outcomes.

Coleman (1986) discussed the good school district in the context of three major descriptions of instructionally effective schools: Goodlad's *A Place Called School* (1984), Lightfoot's *The Good High School* (1983), and Rutter,

Maughan, Mortimore, Ouston, and Smith's *Fifteen Thousand Hours* (1979). Common features of the three studies included the focus on "culture" as opposed to resources; the consequent emphasis on "thick description"; and the holistic concern that syndromes, configurations, and patterns are important, rather than single variables such as class size. The good school has a dual focus: It pursues concurrently academic (focusing on learning, accountability, and change) and nurturance (focusing on caring, collegiality, and community) purposes (Coleman, 1986, p. 90). Coleman (1986) concluded his discussion by suggesting that good school districts have characteristic norms and practices, labeled ethos, which have classroom, school, and district level consequences. The linkage among these elements is coordination (p. 95).

Purkey and Smith (1983) found some commonalities in their review of effective schools: strong leadership by the principal and other staff; high expectations by staff for student achievement; a clear set of goals and an emphasis for the school; an effective schoolwide staff training program; and a system for monitoring student progress (p. 435). The authors proposed nine organization-structure variables as important: school-site management; instructional leadership; staff stability; curriculum articulation and organization; schoolwide staff development; parental involvement and support; schoolwide recognition of academic success; maximized learning time; and district support (pp. 443-445). They recommended longitudinal studies in a variety of schools that track school and student performance over time; a fuller investigation of the process by which schools increase, decrease or maintain effectiveness; identifying information about the procedures employed, the obstacles encountered, and the intended and unintended results obtained; and finally, identification of the intermediate steps of "goals specification and problem diagnosis" (pp. 447-448).

Oakes (1989) identified three global school conditions that act as enablers for high-quality teaching and learning: access to knowledge, press for achievement, and professional teaching conditions. Mediating factors include school resources, alterable organizational policies and structures, and school culture. Some alterable policies and structures are allocation of time (for instruction); curricular emphasis; grouping practices (tracking based on abilities); extra support for low achieving students; and amount and kind of parental involvement (especially at the elementary level). Aspects of school culture that are conducive to quality are commitment to student learning; primacy of teaching; collaborative staff planning/sharing and teamwork; and opportunities for program improvement and professional renewal. Global school conditions function synergistically—they act as enablers rather than as causes of quality teaching and learning.

The OECD international report on schools and quality (1989b) concluded that "student motivation and achievement are profoundly affected by the distinctive culture or ethos that is to be found in each school" (p. 141). The report identified ten characteristics that appear to play a role in determining school outcomes. Some are similar to those identified by Oakes (1989), such as collaborative planning; staff development; and parental involvement); others include the commitment to clearly and commonly identified norms and goals; leadership in initiating and maintaining improvement; staff

stability; maximum use of learning time and the active and substantial support of the responsible education authority (p. 140).

Impact on Students

The implementation of an indicator system should deal with the issue of individual differences in students, the most prominent of which is differences in ability levels. The use of tracking or streaming—students in different programs/courses based on prior achievement—has been the most common method of accommodating individual differences. Although tracking may be beneficial for some groups of students in the high tracks, it can be an ineffective and unfair way of dealing with individual differences because of its damaging effects on students in the low tracks (Oakes, 1987). Oakes identified some of the effects of tracking and examined some of the school and societal contexts for the continued use of tracking despite its negative impacts. She suggested that future research on whether tracking works should consider how, for whom, and toward what ends. In an Ontario study of the relevance of education and the issue of dropouts, Radwanski (1987) also recommended the elimination of streaming.

Nagy, Traub, and MacRury (1986) discussed in detail the impact of proposed province-wide testing in Ontario of selected high school courses. Some of the effects on students are test anxiety, fear of failure, increased number of students entering the general stream, and increased number of students dropping out of the system. The impacts on the teaching-learning processes are intrusion of the exams into the classroom; less emphasis on discovery oriented learning; and more emphasis on strategies to raise test scores. The impacts on institutional policies are the use of scores as accountability tools for schools; and use of scores as an admission criterion to postsecondary institutions.

Administrative and Technical Issues

Many administrative and technical issues are involved in implementing indicator systems. Included in this discussion are strategic considerations, resources, system design, reliability and validity, and data bases.

Strategic Considerations

Currently, different jurisdictions (local, provincial, national) collect a large volume of information (such as annual education statistics, census data, achievement and ability test results, surveys, and polls). The problem is that the information is collected, analyzed, and reported in different ways so that it is difficult to compare and integrate the information from the different sources.

When an indicator system is contemplated, four strategic questions dominate the discussion:

1. How can the information be collected and coordinated?
2. How can it be used effectively and efficiently?
3. Who is responsible for the management of the information?
4. Who is responsible for maintaining the currency of the information?

When information is collected by more than one level of government and/or school jurisdiction, there must be a way to aggregate and disaggregate

the information. Different levels of the educational system require information at different levels of specificity. Information collected at the individual or classroom level by the teacher likely needs to be aggregated to at least the school level to be useful to the district. Nagy (1990) distinguished between differing information needs at the local and provincial levels. What may be important to local authorities intent on improvement at the micro level, might not be the same as that considered by provincial authorities intent on improvement at the macro level. Summary data serving political and accountability purposes may not provide diagnostic information that the school can use for prescribing action. Bock and Mislevy (1988) recommended the duplex design as a way to accommodate differing information needs and to tailor information reports to the specific needs of different users such as students, schools, districts, government, and the media. McEwen and Zatko (1989) also described differing provincial and local information needs. Provincial governments are concerned with achieving a standard that is comparable to national and international counterparts. School jurisdictions are concerned with comparing and interpreting results for their own students in different schools as well as comparing their results with those of other jurisdictions of comparable size, location, and economic base. The two perspectives require some common indicators for provincial purposes and others for local or regional needs.

The implementation of an indicator system may mean developing a way to incorporate existing information or modifying collection procedures to provide integrated information. For example, Fetler (1989) used the California State Department of Education data base to investigate factors that contribute to the dropout rate. These included total enrollment, enrollment in academic courses, achievement scores, and index of poverty. The study illustrates the conceptualization of a model for predicting dropout rates utilizing existing data bases. The most difficult and crucial issues are the conceptualization of the problem and the selection of appropriate variables.

An important strategic consideration is coordination (Luke, 1986; McEwen, 1990; Zatko, 1990). People responsible for indicator systems need to have considerable technical expertise to appreciate the complexities of research design and implementation (such as appropriate methodologies, reliability, validity, etc). They also need highly developed administrative skills to deal with the management, coordination, and dissemination of information to appropriate audiences. Moreover, they need persuasive skills to negotiate the desires and expectations of various stakeholders (administrators, teachers, parents, the public, others) and to be able to communicate clearly and effectively with different groups (government, school personnel, research community, the public). People who possess a combination of all these skills are rare, yet they are essential to ensure that the indicator system can capitalize on the various levels of expertise provided by the different groups.

Resources

Implementing an indicator system requires the commitment of considerable resources: information, staff, money, and time. Kanter (1989) summarized the issue of resources by suggesting capitalizing on existing strengths and

tackling opportunities while living within constraints. The challenge of doing more with less is the ultimate balancing act (p. 372).

McDonnell and Elmore (1987) described three types of information to assist in shaping the choice of policy and/or indicators: political (what is preferred by other policy makers, organized interest groups, and constituents); strategic (information about the target, its capacity to implement, and its probable response to various policy instruments); and analytical (information about the technical requirements of various instruments and which ones are likely to work under different conditions). The last two types of information are particularly important for policy areas such as education, where control is fragmented among different actors and levels of government (p. 25). Oakes (1989) concurred with the importance of information about resources, policies, organizational structures, and processes that characterize schools. In recognition of the importance of informed discussion, both OECD and Alberta Education are providing technical documents to support their respective indicator projects. For example, the EQI initiative has produced an annotated bibliography, a methodological report, and an inventory of assessment instruments.

Staff commitment and expertise are also important ingredients for the successful implementation of an indicator system. Without shared expectations and the support of the community in the determination of the goals, priorities, and expectations for its students, a local school jurisdiction may not address the questions its constituents want answered. Lack of involvement in and development of commitment to such an undertaking may lead to attempts to thwart the system. Odden and Marsh (1988) recommended collaboration and initial involvement of key players to build commitment. They also recommended using cross-role teams, training and assistance, and continuing leadership support and pressure as ways to keep people interested in and committed to projects.

Financial assistance is necessary for the implementation of any innovation. In times of economic restraint, however, it usually means reallocating existing resources rather than finding new sources of revenue. One way to deal with the financial issue is to allocate resources in kind, for example, releasing staff members from current obligations to undertake new tasks, reassigning priorities for existing staff and financial commitments, and eliminating redundant or unproductive activities.

The importance of time cannot be underestimated. Assessing students takes considerable time; indicators must be selected judiciously so that the minimum amount of instructional time is required for measuring student outcomes. Sampling strategies such as those proposed by Maguire and Rogers (1989) and McLean (1987) can help to minimize the amount of time any individual student needs to be involved. Collecting information about the conditions (context, inputs, processes) may not disrupt instructional time but can exact a toll in terms of collecting, analyzing, and reporting results. Deploying more staff to complete these tasks can reduce elapsed time but may result in less accuracy because of the greater number of people working on different parts. Competing demands of accuracy and timeliness become a major issue.

System Design

The implementation of an indicator system presents many technical issues such as sampling, data collection, generalizability, reliability, and validity. Two research projects exemplify the differences between a centralized and a decentralized mode of operation. Frymier, Barber, Gansneder, and Robertson (1989) detailed some of the techniques used in a collaborative research project of students at risk with 276 schools in 87 Phi Delta Kappa (PDK) chapters. Data were collected from over 22,000 students and over 9,000 teachers. Reliability of the data collection was increased by having a small group of trained coders code the data from all sites. Analyses were also centralized, using the same statistical procedures and programs. This study illustrates a practical methodology for research with organizations that have both centralized and decentralized features. The PDK study can be contrasted to the approach and the resulting issues identified by the Canadian research team involved in the Second International Science Study (Connelly, Crocker, & Kass, 1988). Data were collected from over 20,000 grades 5, 9, and 12/13 students and their teachers. Canada was divided into three regions: western (the territories and the western provinces); central (Ontario); and eastern (Atlantic provinces). Within each region, the project was coordinated by a principal research team at a university. In effect, the project was akin to three separate but simultaneous projects because each of the coordinators had to conduct his or her portion within the jurisdiction of his or her university. Unlike the PDK study, there was no central organizing and coordinating structure.

Indicator systems should be as flexible as possible to provide for planning, policy and decision making purposes. This implies that improvement is the goal but that accountability must be demonstrated. There is no single appropriate indicator system, since each must be tailored to the needs, goals, priorities, and expectations of its originators. Thus, while a provincial government might be concerned about determining the efficiency and effectiveness of its educational system based on credible criteria for assessing performance, the federal government might be interested in the comparison of Canadian students with their international counterparts on valued knowledge and skills.

Reliability and Validity

For the establishment of any system of indicators, the mechanisms or tools for measuring the indicators must be reliable and valid. Although reliability of the instruments used to measure the indicators is always important, validity is paramount. That is, one should not steer away from potential indicators that are meaningful (high validity) but difficult to measure (low reliability). The current discussions on the need for more authentic student assessment of overall achievement aim to develop measures of highly meaningful skills, such as critical thinking, in more reliable ways (Frederiksen & Collins, 1989; Norris, 1989; Wiggins, 1989). This is contrasted with the current reliance on easy-to-score, multiple-choice questions that emphasize lower order thinking skills. McLean (1987) noted that many of the well constructed, psychometrically sound tests have strong validity for purposes

of sorting and ordering students but have low pedagogical validity because the tests do not suggest ways to improve learning. Ideal measurements should reflect school learning (the traditional notion of validity) and suggest improvements (pedagogical validity). McLean called pedagogical validity the soul of assessment and accountability its body or visible manifestation (p. 4).

Messick's (1988) distinction between evidential and consequential validity of an assessment instrument is also relevant for indicator systems. He argued that "the process of test use inevitably places test scores in both a theoretical context [evidential basis] of implied relevance and utility and a value context [consequential basis] of implied means and ends" (p. 41). Test developers and users have to be critical of these aspects of test use and interpretation when evaluating validity. There are clearly defined guidelines regarding the use and dissemination of test results (e.g., American Psychological Association, 1986; Code of Fair Testing Practices in Education, 1988; ETS Standards for Quality and Fairness, 1987; Standards for Educational and Psychological Testing, 1985). This is not the case with indicators such as demographics, student-teacher ratios, salaries, and so forth. Their nontechnical nature may lead to misuse and misinterpretation because of their deceptive simplicity. The actual meaning or interpretation of these types of indicators may be extremely complex.

Data Bases

An essential requirement for developing and implementing an indicator system is the establishment of a data base. Aside from issues of reliability and validity, one criterion for judging a data base is its currency. Currency of a data base should not be confused with the use of a data base to track yearly or periodic trends. The issue of currency for standardized tests relates to judging individual performance against the normative sample provided by the test developer. Koretz (1988) cited the lack of currency of the norms of standardized tests as one of the primary reasons for the Lake Wobegon effect (where many states performed above average). Because renorming a standardized test is an expensive and time-consuming process, the test user is forced to make trade-offs between an instrument that has excellent psychometric properties but dated norms and another that may be weaker psychometrically but has more up-to-date norms. One can supplement dated norms by maintaining local norms at the district or provincial level. Another solution is to combine normative information from the developer and to set a standard/criterion for specific schools/districts. Currency is also a concern for nontest score data bases. Once a commitment is made to establish a data base, it is important to ensure that there are established procedures and mechanisms for updating the information.

Another aspect of establishing and maintaining a data base is the consistency and comparability of the information collected. These issues become more complex as the number of reporting units increases and the number of instruments proliferates. For example, there are several instruments for measuring self-esteem. Similarly, a school system may note that nutrition is an important indicator of physical health, but individual schools may use different proxy measures of nutrition. Comparability is also an issue when measures evolve and change over time. Test developers in the cognitive

domain have developed techniques (such as vertical and horizontal scaling) for dealing with comparability. Techniques to deal with comparability for instruments in the affective and behavioral domains have generally been less sophisticated (such as correlation coefficients). Frymier et al. (1989) essentially dealt with this problem by insisting that individual PDK chapters use the same set of measures for the predetermined indicators/variables and allowing flexibility in others.

Political and Financial Issues

Political and financial issues are probably the most complex and difficult to resolve. Negotiation and compromise are required to deal with the many interrelated aspects of society. There are no solutions, only political trade-offs.

Competing Priorities

Education today is one of the many competing public policy areas. It is not independent of the social and economic aspects of society. Luke (1986) identified complexity and uncertainty as major issues. Globalization is another. Technology has lessened geographic and social distances so that the world is a smaller place where interaction of international, national, and local systems become more pervasive, intense, and visible (p. 75). Austrom and Lad (1989) characterized the turbulent environment as one of rapid change and uncertainty, which is a function of the complex and unpredictable changes in the social and geopolitical environment. As the rate of change has accelerated, the issues confronting organizations have emerged as interdependent problems (p. 23). Fiscal restraint, coupled with changing demographics and changed demands, has given rise to a variety of social needs that have not been addressed adequately through traditional institutional mechanisms (p. 28). Efficiency and effectiveness have become two criteria that dominate the public policy agenda.

Current trends in public policy include higher goals for society in economic, social, political, and science policy; increased benefits for both privileged and less privileged groups; the use of positive incentives for encouraging socially desired behavior; a more pragmatic, mixed approach in dividing responsibility between the public and private sectors; and the use of multicriteria decision making (Nagel, 1988, pp. 13-14). Hunter (1988) identified 13 multifaceted and controversial issues that have and will continue to dominate the public policy agenda. Among these are education and training for a quality life and work force; refocusing on families and communities; and human and individual rights and values in practice. He suggested that a leadership challenge is to refocus on families and communities as key institutions in carrying out public policy while maintaining the rights and freedoms of individuals (p. 48). Among the strategies he proposed for managing change are recognizing the global scope of issues, the long-term nature of issues, and the crossing of traditional boundaries (p. 55). The current leadership challenge is to avoid crises while making changes to reduce conflict and restore balance; that is, to inform and engage people in building their own future by addressing the conflicts and imbalances in society and the economy and with other nations openly and honestly (p. 57). Luke (1986) suggested that the

interconnected context requires a "catalytic leader," that is, one who engages in strategic thinking and facilitates the development of a shared vision and collaborative goal (p. 81).

In this context of competing priorities and fiscal restraint, educational authorities (local, provincial, or national) establish priorities in order to allocate scarce resources. For example, the 1988 Alberta School Act established five principles to ensure the fundamental purpose of education to ensure that students learn. These principles are access to quality education, equity, flexibility, responsiveness, and accountability (Alberta Education, 1987, p. 4). Alberta's Minister of Education added three priorities to guide the direction of education: excellence, equity, and efficiency. These are similar to those proposed by Boyer (1989) who identified five priorities for the systemic rather than symbolic push for excellence: goals, equality, teachers, school leadership, and accountability. These two examples illustrate how priorities set the agenda for education within the context of the larger society.

Capacity and Constraints

Implementation of an indicator system needs to consider some of the competing forces at the local level that may either hinder or foster change. Firestone (1989) distinguished two aspects of the school system that influence possible actions toward reform, will and capacity. Will is determined by the responses of the dominant coalition (usually consisting of the superintendent, the principal, and other school staff) in the specific school district toward the proposed reform. The two critical dimensions of the group are propensity for action and the perceived utility of the reform or policy. The most active use of reform results from a combination of a high propensity for action and a positive perception of utility of a particular reform or policy. Three elements that are crucial for determining a district's capacity are ease of personnel mobilization; necessary leadership qualities for the implementation of reform (such as providing and selling a vision, obtaining resources, handling disturbances); and positive district-school linkages (such as conveying the attitude that the district is serious about reform, provision of targeted support, ownership of reform by teachers). In order for the implementation to be successful, an indicator system must be perceived as useful, the dominant stakeholder groups should have a high propensity for action, the organizational structure must allow for personnel mobilization, there is strong leadership, and there are positive district-school linkages.

Lam (1985) developed a two-dimensional matrix for measuring the constraints on schools. His conceptual framework can be extrapolated to encompass schooling (government and school jurisdictions). According to his framework, there are four domains that constrain schooling. The political includes policy and control; the economic includes funding and resources; the social includes values and demographics; and the cultural includes ethnic groups and language (pp. 372-374). Five environmental attributes: change, complexity, organization, routineness (certainty), and directness can describe the external factors. This description can provide an understanding of the environment (p. 379). Zatko (1990) also commented on the difficulties in reforming education: the tension among top-down, collaborative, and

hybrid approaches to change; system inertia; the maintenance of sustained reform efforts; complexity and coordination; scope of reform; trust and collegiality; the differences between means and results; the desire for a blueprint versus a general direction; and the importance of communication.

Berry (1989) identified four interacting bases for transforming schools: vision (needs, beliefs, priorities), ownership (power, responsibility, decisions), capacity (role purpose, resources, operations), and support (organization and work structure). For successful change to occur, there needs to be a level of dissatisfaction with the existing order, a clear vision of what is desired, the identification of practical steps to effect change, and a belief that change is possible. These elements must be greater than the cost of change. Kanter (1989) suggested six criteria for forging successful partnerships and alliances over the long term: the relationship is important and gets adequate resources, management attention, and sponsorship; longer-term investment tends to help equalize benefits over time; the partners are interdependent which helps keep power balanced; organizations are integrated so that contact and communication are managed; each is informed about the plans and directions of the other; the partnership is institutionalized through a framework of supporting mechanisms, from legal requirements to social ties to shared values, which make trust possible (p. 173).

Implementation Strategies

This section presents some strategies for minimizing negative impacts in the implementation of indicator systems. Odden and Marsh (1988) developed a conceptual framework for studying education reform implementation research. They recommended that it should integrate analysis of the content of the reform, the process of its implementation in the local setting, and its effects; focus on the influence of the reform on the overall local educational system as well as on the content, implementation process and more specific impacts; integrate a macro (provincial/state level) with a micro (district/school level) focus for analyzing the above issues; draw on the distinction between developmental and redistributive types of government programs; use recent research on the local change process and relate the results to the macro context, to the content of the reform and to the outcomes at the local level; and identify several types of outcomes, including impacts on the individuals within local educational systems and impacts on the systems themselves (p. 45).

Virtually all the reports cited in the earlier sections identified the need to provide multiple perspectives on education in order to ensure that the indicator system will have the support of the key people in making it work. Therefore, the proposed strategy might be called the multiplier effect:

1. multiple goals of education, based on appropriate dimensions and domains of schooling;
2. multiple indicators of each goal measured by multiple methods;
3. multiple levels of analysis: student, class, school, system, province, (and potentially) country, the world; and
4. multiple participants: government, administrators, teachers, academics, parents.

Multiple goals are important because schooling is not a unidimensional activity. Most educational authorities have incorporated the development of student learning in cognitive, affective, behavioral, and social domains into their programs of study. Yet little information exists about students' progress in developing desired concepts, skills, and attitudes. Measuring desired performance beyond the cognitive domain requires more sensitive instruments than either criterion or norm-referenced testing approaches. Suggestions for developing pedagogical value in assessment were mentioned earlier (e.g., Frederiksen & Collins, 1989; McLean, 1987; Wiggins, 1989). In order to interpret the results and propose intervention strategies, it is also necessary to collect information about the conditions (context, inputs, processes) of schooling. Both the OECD and EQI projects have incorporated conditions into their indicator systems.

Multiple indicators of desired goals are important because each method of measurement has inherent philosophical and technical bases (House, 1983; Noblit & Eaker, 1988; Stufflebeam & Webster, 1988). If more than a single method is used to provide information about a particular goal, with its variations, more confidence in the findings will result. It is also easier to decide on improvement targets and strategies if the information comes from multiple sources. If parents, teachers, and administrators all agree that student attendance requires improvement, based on a school's attendance records and anecdotal evidence, the consensus of what the problem is will result in better ways to ameliorate the situation. Applying these indicators to the CIPP model might suggest that low achievement and poor self-esteem (cognitive and affective outcomes) originating in an economically depressed inner city school (context) might be ameliorated by appropriate interventions (process) made by reallocating staff, time, and other resources (inputs).

In addition to using multiple indicators for multiple goals, it is important to collect information so that multiple levels of analysis are possible. This has been referred to as micro versus macro analysis (Cziko, 1989; Odden & Marsh, 1988; Nagy, 1990) or as methods of aggregating data (Bock & Mislevy, 1988; McEwen & Zatko, 1989). Each of the various educational partners (students, parents, teachers, administrators, government, researchers, public) has different needs for information and each has a valuable perspective to contribute in interpreting results and suggesting improvement targets. Embedding multiple levels of analysis into the design of an indicator system will assist the different partners to share in the planning, setting policies, suggesting instructional practices, and making appropriate decisions for their particular constituent groups. Analyzing local and provincial results is possible when common goals, curricula, and assessment strategies are mandated. National and international comparisons introduce greater complexity due to differing educational philosophies, curricula, languages, organization, and so forth.

Finally, the key to the successful implementation of an indicator system is to involve the partners in the process (Luke, 1986; Kanter, 1989). An indicator system will succeed only if it is viewed as valuable and important. While each group of partners has its own agenda for reform and improvement, they can work together to take action on their particular area of

responsibility. Education, like other public policy areas, changes through negotiation and compromise among competing values and expectations. That this compromise should result on the basis of informed opinion rather than through vested interest is paramount. An effective indicator system, the goals, analyses, and methods of which represent multiple perspectives on education, can make a major contribution to improving education and, therefore, the economic and social prospects for the next generation.

Conclusion

The implementation of indicator systems is a complex activity fraught with potential pitfalls. Issues can be categorized as educational and social, administrative and technical, and political and financial. The negative impacts can be minimized if an indicator system is perceived as an integrated whole in which the various aspects function as a system. Multiple indicators are needed to reflect the complexities of schooling. There is no single recipe for an indicator system. Ideally it should be rooted in a conception of the important components of the operation and outcomes of schooling. The system should be based on the goals, priorities, and expectations of its originating group, be it local, provincial, national, or international. The more stakeholder groups that are involved, the greater the need for discussion and agreement as to the most salient and important aspects for monitoring. Successful implementation necessitates understanding the issues so that negative impacts can be minimized. Resolution often involves compromise among competing perceptions and values. The benefits of an effective integrated information system that provides a comprehensive picture of the broad aspects of schooling are worth the risk.

Acknowledgments

We would like to thank Thomas Maguire, Robert Crocker, and two anonymous reviewers for their valuable suggestions to an earlier version that was presented at the 1990 meeting of the Canadian Educational Researchers' Association.

References

Alberta Education. (1987). *Framework for a new School Act. Highlights*. Edmonton: Government of Alberta.

American Psychological Association. (1986). *Guidelines for computer-based tests and interpretations*. Washington, DC: Author.

Austrom, D.R., & Lad, L. (1989). Collaborative approaches to turbulence in the sociopolitical environment: The emergence of issues management alliances. *Futures Research Quarterly, 5*(4), 22-42.

Berry, S. (1990, February). *Collaborative structure for outcome-based education*. Paper presented at the annual conference on outcome-based education, Phoenix, AZ.

Bock, R.D., & Mislevy, R.J. (1988). Comprehensive educational assessment for the states: The duplex design. *Educational Evaluation and Policy Analysis, 10*(2), 89-105.

Boyer, E.L. (1989, October 1). School reform. A national strategy. *Vital Speeches of the Day,* 741-744.

Brandt, R. (1989). On misuse of testing: A conversation with George Madaus. *Educational Leadership, 46*(7), 26-30.

Carroll, J.B. (1963). A model of school learning. *Teachers College Record, 64*(8), 723-733.

Code of fair testing practices in education. (1988). Washington, DC: Joint Committee on Testing Practices.

Coleman, P. (1986). The good school district: A critical examination of the adequacy of student achievement and per pupil expenditures as measures of school district effectiveness. *Journal of Education Finance, 12*(1), 71-96.

Connelly, F.M., Crocker, R.K., & Kass, H. (1985). *Science education in Canada. Volume 1: Policies, practices, and perceptions*. Toronto, ON: OISE Press.

Connelly, F.M., Crocker, R.K., & Kass, H. (1988). National curriculum research: Problems in the metamethodology of studies concerned with state policy and local variation. *Comparative Education Review, 32*(4), 430-451.

Connelly, F.M., Crocker, R.K., & Kass, H. (1989). *Science education in Canada. Volume 2: Achievement and its correlates*. Toronto, ON: OISE Press.

Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher, 19*(1), 3-13.

Cziko, G.A. (1989). Unpredictability and indeterminism in human behavior: Arguments and implications for educational research. *Educational Researcher, 18*(3), 17-25.

ETS Standards for quality and fairness. (1987). Princeton, NJ: Educational Testing Service.

Fetler, M. (1989). School dropout rates, academic performance, size, and poverty: Correlates of educational reform. *Educational Evaluation and Policy Analysis, 11*(2), 109-116.

Finn, C.E., Jr. (1988). Education reform in the United States: Trends and issues. *International Journal of Education Research, 12*, 115-121.

Firestone, W.A. (1989). Using reform: Conceptualizing district initiative. *Educational Evaluation and Policy Analysis, 11*(2), 151-164.

Fraser, B.J., Walberg, H.J., Welch, W.W., & Hattie, J.A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research, 11*(2), 147-252.

Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*(9), 27-32.

Frymier, J., Barber, L., Gansneder, B., & Robertson, N. (1989). Simultaneous replication: A technique for large-scale research. *Phi Delta Kappan, 71*, 228-231.

Goodlad, J.I. (1984). *A place called school*. New York: McGraw-Hill.

Hargreaves, A. (1989). *Curriculum and assessment reform*. Toronto, ON: OISE Press.

Health and Welfare Canada. (1989). *Charting Canada's future: A report of the demographic review*. Ottawa: Minister of Supply and Services Canada.

House, E.R. (1983). Assumptions underlying evaluation models. In G.F. Madaus, M. Scriven, & D.L. Stufflebeam (Eds.), *Evaluation models: viewpoints on educational and human services evaluation*. Hingham, MA: Kluwer-Nijhoff.

Hunter, K.W. (1988). Facing the public policy issues in 1989. *Futures Research Quarterly, 4*(3), 37-58.

Hymel, G.M. (1988, April). *A macromodel of effective, outcome-based mastery learning school variables: An expanded view*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Kanter, R.M. (1989). *When giants learn to dance: Mastering the challenge of strategy, management, and careers in the 1990s*. Toronto, ON: Simon and Schuster.

Koretz, D. (1988). Arriving in Lake Wobegon: Are standardized tests exaggerating achievement and distorting instruction? *American Educator, 12*(2), 8-15.

Koretz, D. (1989). The new national assessment: What it can and cannot do. *NEA Today, Special edition—Issues '89*, 23-37.

Lam, Y.L.J. (1985). Toward the construction of a school environment instrument: A conceptual framework. *Canadian Journal of Education, 10*, 362-382.

Lightfoot, S. (1983). *The good high school*. New York: Basic Books.

Luke, J.S. (1986). Managing interconnectedness: The need for catalytic leadership. *Futures Research Quarterly, 2*(2), 73-83.

Maguire, T.O., & Rogers, W.T. (1989). Proposed solutions for nonrandomness in educational research. *Canadian Journal of Education, 14*, 170-181.

McDonnell, L.M., & Elmore, R.F. (1987). *Alternative policy instruments*. Washington, DC: Center for Policy Research in Education.

McEwen, N. (1990, April). *Educational quality indicators: Developing indicator systems in Alberta*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

McEwen, N., & Zatko, G. (1989, June). *Educational quality indicators: Framework for action*. Paper presented at the annual meeting of the Canadian Educational Researchers' Association, Quebec City.

McLean, L.D. (1985). *The craft of student evaluation in Canada*. Toronto, ON: Canadian Education Association.

McLean, L.D. (1987, April). *Emerging with honour from a dilemma inherent in the validation of educational achievement measures*. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Meisels, S. (1989). High-stakes testing in kindergarten. *Educational Leadership, 46*(7), 16-22.

Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H.I. Braun (Eds.), *Test validity* (pp. 33-45). Hillsdale, NJ: Lawrence Erlbaum.

Nagel, S.S. (1988). Trends in public policy in America. *Futures Research Quarterly, 4*(3), 5-16.

Nagy, P. (1990, May). *Must we examine outcomes?* Paper presented at the annual conference of the Canadian Evaluation Society, Toronto.

Nagy, P., Drost, D., & Banfield, H. (1985). A path analytic study of the influences of community isolation on student achievement. *Alberta Journal of Educational Research, 31*, 209-227.

Nagy, P., Traub, R.E., & MacRury, K. (1986). *Strategies for evaluating the impact of province-wide testing*. Toronto, ON: Ontario Institute for Studies in Education.

Neill, D.M., & Medina, N.J. (1989). Standardized testing: Harmful to educational health. *Phi Delta Kappan, 70*, 688-697.

Noblit, G.W., & Eaker, D.J. (1988). Evaluation design as political strategies. In J. Hannaway & R. Crowson (Eds.), *The politics of reforming school administration: The 1988 yearbook of the Politics of Education Association* (pp. 127-138). Philadelphia, PA: Falmer.

Norris, S.P. (1989). Can we test validly for critical thinking? *Educational Researcher, 18*(9), 21-26.

Nuttall, D.L. (1990, April). *The functions and limitations of international education indicators*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Oakes, J. (1986). *Educational indicators: A guide for policymakers*. New Brunswick, NJ: Center for Policy Research in Education, Rutgers University.

Oakes, J. (1987). Tracking in secondary schools: A contextual perspective. *Educational Psychology, 22*, 129-153.

Oakes, J. (1989). What educational indicators? The case for assessing the school context. *Educational Evaluation and Policy Analysis, 11*(2), 181-199.

Odden, A. (1990). Educational indicators in the United States: The need for analysis. *Educational Researcher, 19*(5), 24-29.

Odden, A., & Marsh, D. (1988). State education reform implementation. In J. Hannaway & R. Crowson (Eds.), *The politics of reforming school administration: The 1988 yearbook of the Politics of Education Association* (pp. 41-60). Philadelphia, PA: Falmer.

Office of Educational Research and Improvement (OERI). (1988). *Creating responsible and responsive accountability systems*. Washington, DC: US Office of Educational Research and Improvement.

Organisation for Economic Cooperation and Development (OECD). (1989a). *Project on International Educational Indicators (INES Project): Review of preliminary results and proposal for future work*. Paris: Author.

Organisation for Economic Cooperation and Development (OECD). (1989b). *Schools and quality. An international report*. Paris: OECD Publications Service.

Purkey, S.C., & Smith, M.S. (1983). Effective schools: A review. *The Elementary School Journal, 83*(4), 427-452.

Radwanski, G. (1987). *Ontario study of the relevance of education, and the issue of dropouts*. Toronto, ON: Ontario Ministry of Education.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge, MA: Harvard University Press.

Selden, R.W. (1988). Missing data: A progress report from the states. *Phi Delta Kappan, 69*, 492-494.

Shavelson, R.J. (1988). Contributions of educational research to policy and practice: Constructing, challenging, changing cognition. *Educational Researcher, 17*(7), 4-11, 22.

Shavelson, R.J., McDonnell, L., Oakes, J., Carey, N., & Picus, L. (1987). *Indicator systems for monitoring mathematics and science education*. Santa Monica, CA: RAND.

Shepard, L.A. (1989). Why we need better assessments. *Educational Leadership, 46*(7), 4-9.

Standards for educational and psychological testing. (1985). Washington, DC: American Psychological Association.

Stufflebeam, D., & Webster, W.J. (1988). Evaluation as an administrative function. In N.J. Boyan (Ed.), *Handbook of research on educational administration* (pp. 569-602). New York: Longman.

Tyler, R.W. (1984). A guide to educational trouble-shooting. *Educational Leadership, 41*(8), 27-30.

Walberg, H.J. (1984). Improving the productivity of America's schools. *Educational Leadership, 41*(8), 19-27.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan, 70*(9), 703-713.

Zatko, G. (1990, April). *The Alberta strategy for educational reform: Balancing inputs, processes, and results*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 1, March 1991, 87-97*

KATHLEEN OBERLE
University of Alberta

Paradigm Wars: Who's Fighting; Who's Winning?

Over the past decade, the qualitative/quantitative paradigm war has continued in educational research. A careful analysis of the epistemological, ontological, and methodological bases for the argument reveals no substantive reasons for clinging to the belief that the qualitative and quantitative paradigms are incompatible. Indeed, most of the participants in the debate elect for a combination of approaches to research. Nonetheless, despite the fact that there is evidence that qualitative research has an important role in the advancement of knowledge in education and psychology, positivist (quantitative) research is still dominant in the literature.

In the past decade, one of the most hotly debated issues in the area of educational research has been whether qualitative methods, quantitative methods, or a combination of both are most appropriate. According to Shulman (1981), "what distinguishes disciplines from one another is the manner in which they formulate their questions ... and the principles of discovery and verification that constitute the ground rules for creating and testing knowledge in their fields" (p. 6). Hence discussions about the nature of science and the appropriateness of research paradigms should be important to researchers in education.

A close examination of developments in educational research reveals that views about science in this field have been influenced largely by debate in other social sciences, particularly psychology, where paradigmatic struggles are similar to those in education. Indeed, many of the questions about "discovery and verification" in education parallel those in psychological research. Moreover, many of the questions of interest to psychologists are also of interest to educational researchers. Therefore, to strengthen the analysis, this discussion of methods appropriate for research in education includes an examination of developments in psychology.

The purpose of this paper is to clarify the current debate. The primary intent of the debate is ostensibly to improve the quality of research by providing researchers with a solid basis for decision making regarding research design. However, a bewildering array of opposing viewpoints concerning the "proper way to do science" are presented in the literature. Faced with a multiplicity of conflicting and often confusing perspectives, the researcher is likely to find the literature singularly unenlightening.

Kathleen Oberle is a doctoral student in educational psychology (measurement) and a registered nurse. Her interest in the qualitative/quantitative debate stems from her work on hospital research review committees.

To illuminate the argument, viewpoints presented over the last decade in two influential journals, *Educational Researcher* and *American Psychologist*, are summarized. Central tenets of the quantitative and qualitative paradigms are first presented. Then, as philosophical discussions about proper science generally contain elements of epistemology, ontology and methodology, arguments in support of the two paradigms are divided into these three domains. The impact of the debate on the actual conduct of research is examined, as evidenced by reported research. The paper concludes with a discussion about the future of the debate and directions for research in education.

Competing Paradigms

Beliefs about the way science should be conducted generally fall under two paradigms. One paradigm adheres closely to positivism and is usually termed quantitative or empiricist; the other paradigm is clearly anti-positivist and is frequently termed termed qualitative or interpretivist/contextualist. Firestone (1987), Manicas and Secord (1983), and Phillips (1983) summarize the positivist position. According to these authors, positivist philosophy demands that scientific propositions be based on observable data. It is assumed that there is correspondence between what is observed and what is "real," The goal of science is to reveal "truth," that is, the lawful relationships between observables. The scientist makes every effort to ensure objectivity of observations. Working from an assumption that a single, explainable reality exists, positivists strive to produce results that are reliable and statistically generalizable. They look for regularities and causal relationships among phenomena. The context in which the observed phenomena occur is considered to be important only insofar as it "contaminates" the data; thus attempts are made to eliminate contextual factors, to control them, or to incorporate them through the research design or statistical procedures. Positivism is generally associated with experimental and quasi-experimental methods.

The positivist paradigm has been dominant in western philosophy until recently and has provided the basis for most scientific endeavor in both the natural and human sciences. However, as Smith (1983) and Phillips (1983) point out, there are those (antipositivists) who challenge the precepts of positivism, particularly for study in the human sciences. Antipositivists maintain that there is not one immutable reality, but rather that reality exists in the mind of each individual. Because human understandings are intertwined with intentions, goals, and purposes, study in the human sciences must incorporate context, rather than attempt to exclude or remove it. Thus the investigator and the subject of investigation cannot be separated. The goal of science, according to the antipositivists, is to search for understanding and common meaning, rather than prediction and control (Gage, 1989). Reductionist and decontextualist methods are deemed inappropriate for study in human science (Jennings, 1986; Kukla, 1986). Various schools of thought fall under the rubric of antipositivism, including contextualism, social constructionism, deconstructionism, critical theory, phenomenology,

and hermeneutics. Associated methods include ethnography, case studies, in-depth interviews, and participant observation.

Philosophical Considerations

Arguments in the qualitative/quantitative debate are often confusing because assumptions and presuppositions underlying the author's position have not been made clear. Moreover, similar arguments may sound different depending on whether the author is arguing from an epistemological, ontological, or methodological perspective. To help clarify the discussion, differences between the three philosophical perspectives are outlined below, and papers representing the various positions are reviewed.

Epistemological Concerns

Epistemology is the study of knowledge and how that knowledge is obtained. The epistemologist is interested in the nature of science and asks questions such as: How should science be conducted? How can progress in science be measured? What are the limitations of scientific knowledge? Also included in the category of epistemology are questions about the role of values in science: Is science value neutral? Is it possible for the scientist to make value free decisions or judgments? Can observation provide unassailable knowledge, or is all knowledge and observation value laden?

In the context of the current debate the central epistemological questions are what science is and what it is not, and how scientific progress should be measured. Until the early 1960s, science was dominated by positivist philosophy, which held that scientific advance was an orderly progression based on the results of carefully controlled experiments designed to provide verification or falsification of hypotheses. Conclusions were deemed to be value free and based on logical deduction, and progress was measured by closer and closer approximations to "truth."

This position was challenged by Kuhn (1970) whose radical ideas provided the basis for much of the current debate. Kuhn's work is interpreted by Manicas and Secord (1983) as suggesting that scientific decisions are not as rule governed as positivists would insist. Instead, Kuhn depicts science as a social activity, in which all observations are value laden. He observes that in each discipline, a set of fundamental beliefs, termed a paradigm, determines what questions the scientist finds interesting and what answers are seen as acceptable. Thus, according to Gholson and Barker (1985), Kuhn suggests that differing paradigms are mutually exclusive. Progress in science occurs when a sufficient number of researchers become dissatisfied with the answers provided by one paradigm and make a radical shift to another perspective. In effect, one paradigm is overthrown by another, and a new set of beliefs forms the basis for decision making. Consequently, incompatibility of paradigms is essential to progress.

Kuhnian philosophy was welcomed by antipositivists who were seeking a way to "legitimize" their endeavor (Phillips, 1983). In the antipositivst view, Kuhn's ideas provided testimony that scientific knowledge is in fact context dependent, and that no single reality exists. Furthermore, Kuhn said nothing about method, and how scientific knowledge should be obtained, which suggested to antipositivists that logical-deductive, positivist techniques were no

longer the only methods acceptable. Mainstream scientists were, however, upset by these suggestions, which to them smacked of relativism. If there were no one reality, and no one way of conducting science, how was one to distinguish between scientific and nonscientific activity? If, as Phillips (1983) suggested, "to be regarded as knowledge, the belief in question has to be justified in some way" (p. 11), how was one to measure the adequacy of justification? Moreover, according to Smith (1983), Kuhnian thought indicated that paradigms were in competition with one another, and therefore incommensurable, which suggested that if positivism was dead, relativism could take over. Fearing that this would happen, positivists lined up in decided opposition to the new philosophy.

There was no doubt that positivism had been seriously challenged, but what was to take its place? Was there to be a new definition of science? Philosophers, including Lakatos (1978) and Laudan (1977) strove to provide alternative epistemologies, somewhere between realism and relativism, that would answer the concerns of the positivists while recognizing the role of values and social factors in science. According to Gholson and Barker (1985), Lakatos acknowledged the existence of certain "hard core" assumptions that formed the basis of what he termed a research program. These assumptions were socially agreed on and essentially unprovable, but, as with Kuhn's paradigm, they formed the foundation for decision making. Scientific progress was evidenced in theories that led to new and better predictions. Lakatos believed that in any one field there could be more than one set of "hard core" assumptions, and that there was no one method that would guarantee scientific advance.

Similarly, Laudan (1977) talked of a research tradition, a "family of theories sharing a common ontology and methodology" (Gholson & Barker, 1985, p. 761), which could change as the research tradition evolved. Progress in science was, to Laudan, the ability of the science to solve problems of interest. Again, as Macmillan and Garrison (1984) suggest, Laudan did not discuss method as such; his philosophy would suggest that a plurality of methods would be acceptable.

How, then, should science be defined in psychology and education? Gholson and Barker (1985) argue convincingly that a blend of positivist and relativist beliefs and methodologies will provide the most satisfactory advance in science in psychology. They base their arguments on the work of Lakatos and Laudan, and maintain that from this perspective a plurality of methods is the only defensible position.

Certainly, positivism alone has failed to provide satisfactory answers to date. As Macmillan and Garrison (1984) note in their criticism of research in education, "a healthy, progressive research tradition should have come up with more significant results than this one has" (p. 18). They suggest that the problem with positivist research rests in the fact that it "is conducted in the language of analysis of variables, but its application is in the intentional language of action and belief" (p. 18). Therefore, based on the epistemologically sound premises of the new philosophy of science, they contend that researchers should use whatever methods they feel will provide them with answers to questions of interest.

Similarly, Mosenthal (1985), in discussing progress in educational research, outlines the definitions of progress put forward by positivists and antipositivists (in this case interpretivists and critical theorists) and submits that all definitions have merit. Citing Lakatos and Laudan, Mosenthal indicates that researchers should have "freedom of choice but not freedom from choice" (p. 7). To him, the important issue is that the researcher make an informed choice, one based on thoughtful understanding of underlying assumptions. Tuthill and Ashton (1983) go as far as to indicate that educational researchers must develop their own paradigm of "common values, assumptions, goals, norms, language, beliefs, and ways of perceiving and understanding the educational process" (p. 9), and that paradigmatic disputes should be solved "democratically," In their view, a plurality of methods is acceptable, provided the researcher recognizes that all scientific knowledge is social knowledge acquired through social processes. These same authors (Tuthill & Ashton, 1984) maintain that the distinction between quantitative and qualitative methods is not an interesting or useful one; what is important is the way in which prior knowledge affects the acquisition of scientific knowledge.

Howe (1985) supports this argument by stating that "rigid epistemological distinctions between quantitative and qualitative methods ... rest on positivistic dogmas, and ... post-positivistic thought eliminates the intractable problem engendered by positivistic epistemology of a forced choice between value-laden/qualitative and value-free/quantitative research methods" (p. 10). In a later article, Howe (1988) repeats his position, finding support in the work of much earlier philosophers, including James and Dewey, proponents of what he terms the pragmatic tradition. He contends that the claim of incompatibility between paradigms is essentially an artifact, and suggests that, "far from being incompatible, then, quantitative and qualitative methods are inextricably intertwined" (p. 12).

From the above arguments it appears that the majority of those engaged in the controversy would contend that the most fruitful position for researchers in human science is somewhere between positivism and relativism, that is, in the creation of a new model which would combine the two opposing paradigms. Despite persuasive arguments, however, not everyone is convinced of the commensurability of paradigms. Smith and Heshusius (1986) maintain that the compatibility thesis merely bows to "what works" without thought to the underlying epistemology. Some are still prepared to assume oppposing positions, as evidenced by Eisner's (1983) and Phillips' (1983) spirited debate in which Phillips calls for a return to some positivist beliefs, and Eisner argues that positivism is, and should remain, dead. Others such as Gergen (1985) argue that positivist assumptions are completely untenable, and have no place in science. Thus the debate continues.

Ontological Concerns

Ontology is the study of what it is "to be." It is often difficult to separate ontology and epistemology because one's ontological beliefs strongly influence one's epistemological position. Nevertheless, a particular epistemology does not necessarily follow from a particular ontology. Confusion arises

when participants in the current debate present an ontological position and make assumptions about epistemology, without explicating these assumptions. Therefore, it is important to make a distinction between epistemology and ontology and to be able to recognize the respective supporting arguments.

Ontological questions of interest in the present discussion can be posed at several levels. First, there is the very general question about the nature of reality. For example, philosophers at the farthest extreme of relativism would suggest that reality exists in name only; names, concepts, and labels are used to structure a reality which does not exist outside the mind of the individual. In contrast, realists would suggest that the social world external to individual cognition is made up of hard, tangible, and relatively immutable structures, that is, the social world exists independent of the individual's perception of it. At another level of ontological questioning is the concept of the character of human nature. Are individuals completely autonomous and free willed (voluntarism), or are man's activities dominated or determined by the environment (determinism)? Because of the close link between epistemology and ontology, the ontological commitment to realism or relativism was implicit in many of the arguments discussed above, and no further analysis of that aspect of the argument will be made here.

Leaving aside realism and relativism, the central ontological question in education is related to beliefs about the determinacy of human behavior. This same question has received intensive scrutiny by psychologists. Cziko (1989) contends that the issue is in fact not whether the individual's behavior is determined, but whether it is predictable. He argues that "complex human behaviour of the type that interests educational researchers is by its nature unpredictable if not indeterminate" (p. 17) and supports his position from both deterministic and indeterministic perspectives. In his view, even if behavior is determined, interaction of the individual with the determinants is so complex that the behavior cannot be predicted. He shows that "even though the relationship between two variables may be both quite simple and completely deterministic, a nonlinear relationship may nevertheless lead to outcomes that are entirely unpredictable" (p. 19). From an indeterministic perspective, he provides evidence for the influence of human consciousness and free will on human behavior. Thus, with regard to educational research, he questions the validity of strictly quantitative, experimental approaches, concluding that, because of the unpredictability of human behavior, research must be limited to description and interpretation.

Kimble (1984), in his examination of value systems of psychologists, finds most respondents to have leanings in the direction of humanism (antipositivism). The one exception is the question of determinacy. He notes that all psychologists are determinists to some degree. Royce (1982) shows agreement with this position in suggesting that the problem of determinacy, which he considers one of the central philosophic issues facing psychology today, "needs to be recast in terms of degrees of determinism, depending on the behaviour in question" (p. 262).

Although psychologists may agree that all behavior is at least to some extent determined, the argument does not end there. Staats (1981) outlines

some important ontological questions that he believes are divisive in psychology. In his conception, there are essentially two psychological camps, the behaviorists and the traditionalists, and the two groups have very different ontological positions. For example, on the question of the intentionality of behavior he notes that "the traditional conception of humans as active, striving, self-directing, cognitively aware, and purposive organisms" (p. 240) is challenged by the behaviorist view of human behavior as mechanistic and passive. He suggests that "psychologists who adhere to one of these two conceptions reject the other and its associated works" (p. 240). Another significant question, in Staats' view, is the concept of personality. Behaviorists, he maintains, believe that there are no internal personality processes, whereas traditionalists believe that the individual has personality structures or processes that determine behavior.

In each area of psychology, particular ontological questions arise and parallel questions are raised in education. For example, in developmental psychology, the concern is "what is the nature of the child?" In an interesting conversation among four of the most prominent developmental psychologists, agreement was reached that "the child, or more specifically, our conceptions of the child are cultural inventions (or constructions)" (Bronfenbrenner, Kessel, Kessen, & White, 1986, p. 1227), and that behavior is intentional and partially determined by person-environment interaction. Clearly such an ontological position would lead the researcher in either education or psychology to consider other than positivist epistemology and methods.

Methodological Concerns

Methodological questions can be discussed at more than one level of abstraction. Smith and Heshusius (1986) make a distinction between "method as technique and method as logic of justification" (p. 4). In considering method as logic of justification, one must attend to the relationships among epistemology, ontology, and method. Discussions that focus on method as technique, on the other hand, are at a lower level of abstraction. Here the concern is with specifics of a particular method, with little regard to the underlying epistemological or ontological assumptions.

From the discussions above, the linkage between epistemology, ontology, and method as logic of justification should be evident. Radical positivists have a realist and determinist ontological perspective; thus they demand objective methods. Radical antipositivists have a relativist and indeterminist position; they demand subjective, interpretive methods. Those who accept a middle-of-the-road epistemology maintain that subjective and objective approaches need not be mutually exclusive; they advocate a blending of methods.

Shulman (1981) points out that it is critical to engage in discussions about proper method, because in any research endeavor, confidence in the method used will determine confidence in the findings. One important methodological question is related to the generalizability of results. Here again, the link with ontology is evident; if one believes in a single reality, generalizability is a central concern. In contrast, if one believes in multiple realities, one must

either refute the importance of generalizability or accept an alternative definition which allows for generalizability of meaning rather than "fact."

The question of generalizability is important in applied disciplines such as psychology and education. Firestone (1987) and Shulman (1981) maintain that if the objective of research is to answer questions of interest, the best approach is the one that supplies the most useful answers, and in an applied discipline answers must be generalizable to more than one situation if they are to inform practice. Each mode of inquiry, says Firestone (1987), has its own strengths. To illustrate, he describes two studies on the same topic but using different methods. The quantitative study "uses a hydraulic image of determinism as if pressure from one variable changes another," whereas the qualitative study "presents a more complex view of a world in which there are limits and opportunities that individuals must take into account and use" (p. 19). The advantage of the quantitative study is that it shows "a pattern that extends across a large number of situations," while the strengths of the qualitative study are a "concrete depiction of detail, portrayal of process in an active mode, and attention to the perspectives of those studied" (p. 20). Both are important, argues Firestone, for in coming at the same problem from different directions, they provide different kinds of information. The quantitative study alone fails to provide the insights and understandings necessary to inform practice; the qualitative study alone fails to provide generalizability. Combining methods affords richer, more useful results.

Not everyone agrees with Firestone's position. Banaji and Crowder (1989) suggest that the present preoccupation with what they term "ecological validity of method," that is, with qualitative studies that seek to provide insight and understanding, leads to research that has limited usefulness. They maintain that "an obsession with ecological validity of method can compromise genuine accomplishments" (p. 1185). In Dar's (1987) view, "we should direct our energy into building integrated research programs capable of producing strong and meaningful predictions and contributing to real scientific progress in psychology" (p. 151). Such research programs do not include qualitative approaches. Similarly, Rosnow and Rosenthal (1989) call for a strengthening of existing quantitative methods, not for new modes for research.

At the level of method as technique, there are numerous arguments in each paradigm about how to improve the calibre of research. For example, Rosnow and Rosenthal (1989), who are clearly positivist in their view, describe four main difficulties in current (positivist) research practice: the overreliance on dichotomous significance testing decisions; the tendency for studies to have low power; the frequent definition of results in terms of significance tests alone; and the failure of scientists to conduct replication studies.

In the qualitative paradigm, Miles and Huberman (1984a, 1984b), who admit to being "close to the middle ground" (1984a, p. 23) in their epistemology, propose a model for analysis of qualitative data. They suggest techniques for data reduction, data display, and conclusion drawing. Marshall (1984) criticizes the model, arguing that Miles and Huberman are trying to force qualitative analysis into positivist modes. She argues that "qualitative re-

search must not be beaten into submission to the approximations of the methods and criteria set by positivists ... It must preserve the chance to explore, capitalize on divergent analyses, and continue to capture the insights of the human research tool" (p. 28).

In summary, in the methodological debate discussions pertaining to method as logic of justification refer to whether subjective, qualitative methods yield useful results, or whether objective, quantitative methods are best. At the level of method as logic of technique, the debate centers on the best way to carry out research in each paradigm.

Impact of the Debate on Educational Research

After more than a decade of debate regarding the nature of proper science in education and psychology, it seems fairly clear that radical positivism is seriously ailing. Kuhn and others have demonstrated that science is not value free; that social forces play an essential role in how observations are made and interpreted; and that progress in science may be defined in many different ways. Philosophers of science have defined a new epistemology that recognizes the role of human interactions in scientific endeavor. In this epistemology, the assumption for human science is of an underlying pattern or structure of human behavior, a pattern that is not immutable, but socially constructed. The goal of science is defined as problem solving, and acceptable methods include any that can help to solve the problems of interest.

Most of those involved in the debate about proper science in education argue in favor of the new epistemology. They suggest that each paradigm can contribute to understanding in education, with qualitative methods used in a mode of discovery, and quantitative methods used for verification and justification. A few cling to the belief that combining paradigmatic assumptions, that is, adopting an epistemology that argues for some degree of relativism and determinism, while accepting the existence of some structure, is impossible. They maintain that paradigms are incommensurable, because the basic assumptions on which they are founded are polar opposites. If one considers only the most extreme positions of positivism and interpretivism, this position is no doubt correct. However, these authors fail to present convincing arguments as to why it is necessary to adopt such radical views. There appears to be no logical reason why a middle-of-the-road epistemology should not be accepted.

Given that argument, and accepting that there has been dissatisfaction among practitioners regarding the effectiveness of research in informing practice (Macmillan & Garrison, 1984; Schubert, 1980) one would expect that major changes in the practice of research in education would have occurred over the last decade.

At the beginning of the decade, it was evident that the vast majority of studies reported used quantitative methods. Willson (1980) reviewed studies reported in the *American Educational Research Journal* (AERJ) from 1969 to 1978 and indicated that only 2% of the total used qualitative techniques. Has the picture changed dramatically? To answer this question, a review of articles published in AERJ from mid-1988 to mid-1989 was undertaken. Of the 27 studies reported, nine (33%) used qualitative techniques, at least to

some degree. This represents a substantial increase. These results raise an interesting question. If qualitative methods meet epistemological criteria for proper science, and if they can provide insights and understandings required for research in an applied discipline, why does the debate continue? Moreover, why are qualitative approaches not more widely accepted and used? The answer may be found in what might be termed sociopolitical considerations. Eisner (1988) suggests that socialization in a particular method affects the way we view science and scientific methods. The dominance of positivism has, until recently, resulted in a limited exposure of fledgling scientists to alternative (interpretivist) strategies. Thus a bias in favor of positivistic definitions of science remains. This bias can have important influences on scientific practice. As Kupfersmid (1988) and Wachtel (1980) have pointed out, in the scientific community rewards are largely in the form of publications and grants awarded. Editors and members of grant committees are themselves prominent scientists, most of whom have been schooled in positivist methods and many of whom hold qualitative approaches suspect. The result is often rejection of articles and grant applications in which qualitative methods are featured. Kupsfersmid (1988) indicates that "researchers interested in increasing the chances of publishing their empirical investigations [must] avoid creative or innovative experimental designs and [must] concentrate efforts on areas that are easy to test and are noncontroversial" (p.635).

The picture may be changing, however. A greater attention to philosophy of science issues is evident in university graduate programs where novice researchers are trained, and it can only be hoped that they are learning to examine and question the assumptions underlying their choice of research method. Such analysis holds promise for the release of researchers from slavish and unthinking devotion to particular approaches. As these new graduates, and their mentors, exert their influence both in academic settings and on editorial boards and grant committees, paradigmatic distinctions may disappear.

The battle is far from over (Gage, 1989), but there are signs of, if not rapprochement, at least detente. The battle will be won when researchers make their decisions based on informed deliberation, not dogma. As the debate continues, hope grows for true advance in research in education.

References

Banaji, M.R., & Crowder, R.G. (1989). The bankruptcy of everyday memory. *American Psychologist, 44*, 1185-1193.

Bronfenbrenner, U., Kessel, F., Kessen, W., & White, S. (1986). Toward a critical social history of developmental psychology. *American Psychologist, 41*, 1218-1230.

Cziko, G.A. (1989). Unpredictability and indeterminism in human behaviour: Arguments and implications for educational research. *Educational Researcher, 18*(3), 17-25.

Dar, R. (1987). Another look at Meehl, Lakatos, and the scientific practices of psychologists. *American Psychologist, 42*, 145-151.

Eisner, E.W. (1983). Anastasia might still be alive, but the monarchy is dead. *Educational Researcher, 12*(5), 13-14, 23-24.

Eisner, E.W. (1988). The primacy of experience and the politics of method. *Educational Researcher, 17*(5), 15-20.

Firestone, W.A. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. *Educational Researcher, 16*(9), 16-20.

Gage, N.L. (1989). The paradigm wars and their aftermath. A "historical" sketch of research on teaching since 1989. *Educational Researcher, 18*(7), 4-10.

Gergen, K.J. (1985). The social constructionist movement in modern psychology. *American Psychologist, 40*, 266-275.

Gholson, B.,& Barker, P. (1985). Kuhn, Lakatos and Laudan. Applications in the history of physics and psychology. *American Psychologist, 40*, 755-769.

Howe, K.R. (1985). Two dogmas of educational research. *Educational Researcher, 14*(8), 10-18.

Howe, K.R. (1988). Against the qualitative-quantitative incompatibility thesis or dogmas die hard. *Educational Researcher, 17*(9), 10-16.

Jennings, J.L. (1986). Husserl revisited. The forgotten distinction between psychology and phenomenology. *American Psychologist, 41*, 1231-1240.

Kimble, G. (1984). Psychology's two cultures. *American Psychologist, 39*, 833-839.

Kuhn, T.S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Kukla, A. (1986). On social constructionism (letter). *American Psychologist, 41*, 480-481.

Kupfersmid, J. (1988). Improving what is published. A model in search of an editor. *American Psychologist, 43*, 635-642.

Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge, UK: Cambridge University Press.

Laudan, L. (1977). *Progress and its problems. Toward a theory of scientific growth*. Berkeley, CA: University of California Press.

Macmillan, C.J.B., & Garrison, J.W. (1984). Using the "new philosophy of science" in criticizing current research traditions in education. *Educational Researcher, 13*(10), 15-21.

Manicas, P.T., & Secord, P.F. (1983). Implications for psychology of the new philosophy of science. *American Psychologist, 38*, 399-413.

Marshall, C. (1984). The wrong time for mechanistics in qualitative research (letter). *Educational Researcher, 13*(9), 26-28.

Miles, M.B., & Huberman, A.M. (1984a). Drawing meaning from qualitative data: Toward a shared craft. *Educational Researcher, 13*(5), 20-30.

Miles, M.B., & Huberman, A.M. (1984b). Rejoinder to Marshall (letter). *Educational Researcher, 13*(9), 28-29.

Mosenthal, P.B. (1985). Defining progress in educational research. *Educational Researcher, 14*(9), 3-9.

Phillips, D.C. (1983). After the wake: Postpositivistic educational thought. *Educational Researcher, 12*(5), 4-12.

Rosnow, R.L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist, 44*, 1276-1284.

Royce, J.R. (1982). Philosophic issues, Division 24, and the future. *American Psychologist, 37*, 258-266.

Schubert, W.H. (1980). Recalibrating educational research: Toward a focus on practice. *Educational Researcher, 9*(1), 17-21.

Shulman, L.S. (1981). Disciplines of inquiry in education: An overview. *Educational Researcher, 10*(6), 5-12.

Smith, J.K. (1983). Quantitative versus qualitative research: An attempt to clarify the issue. *Educational Researcher, 12*(3), 6-13.

Smith, J.K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher, 15*(1), 4-12.

Staats, A.W. (1981). Paradigmatic behaviorism, unified theory, unified theory construction methods, and the Zeitgeist of separatism. *American Psychologist, 36*, 239-256.

Tuthill, D., & Ashton, P. (1983). Improving educational research through the development of educational paradigms. *Educational Researcher, 12*(5), 6-14.

Tuthill, D., & Ashton, P. (1984). Response to Miles and Huberman (letter). *Educational Researcher, 13*(9), 25-26.

Wachtel, P.L. (1980). Investigation and its discontents. Some constraints on progress in psychological research. *American Psychologist, 35*, 399-408.

Willson, V.L. (1980). Research techniques in AERJ articles: 1969-1978. *Educational Researcher, 9*(6), 5-10.

New Education, first published in 1978, is a long-established and respected journal, which provides an international forum for original ideas and investigations on current and future trends in educational theory and practice. The journal devotes its attention to *educational reforms affecting elementary and secondary curricula, teaching methods, curriculum design, development and evaluation, school syllabuses, educational administration, vocational education, aesthetic and health education, moral education, lifelong education, innovations, participation in education, comparative studies in academic achievement, teacher training and in-service education for elementary and secondary teachers.*

The articles are chosen for their originality, *readability and accessibility* to a *wide* audience of teachers, principals, academics, educational administrators, consultants and community education officers.

In order to ensure that articles are of the highest quality all contributions are submitted to at least two independent referees before acceptance for publication.

Contributions are invited (normally no more than 6000 words, in two copies with stamped addressed envelope for return) and should be sent to the Editor, *New Education*, **James Nicholas Publishers, P.O. Box 244, Albert Park, Australia, 3206.**

The journal is published twice a year, in July and December. An annual index and title-page is bound in the December issue.

1989 Volume 11, **ISSN 0156-0905**

Subscription rates (payable in advance): one year US$90.00 Individuals and elementary and secondary schools US$48.00 (for accelerated surface postage and packing add US$5.00 per issue). Orders and remittances should be addressed to: **James Nicholas Publishers, P.O. Box 244, Albert Park, Australia 3206.**

EDITOR: **Rea Zajda**

CONSULTING EDITORS

Professor Basil Bernstein, *University of London*
Professor Martin Carnoy, *Stanford University*
Professor John Eggleston, *University of Warwick*
Professor Douglas Sloan, *Teachers College, Columbia University*
Professor Arthur Wirth, *Washington University*

RECENT AND FORTHCOMING ISSUES

Lawrence Saha **SOCIAL MOBILITY AND SOCIAL REPRODUCTION**
Martin Carnoy **EDUCATION AND THEORIES OF THE STATE**
Rob White **EDUCATION IN NICARAGUA**
Henry Giroux **EDUCATIONAL REFORMS AND THE POLITICS OF TEACHER EMPOWERMENT**
Ivan Illich **ECO-PAEDAGOGICS AND THE COMMONS**
Elliot Eisner **THE ART AND CRAFT OF TEACHING**
Margaret Sutherland **DEVELOPING CREATIVE ABILITIES OF TEACHERS**
Gary Bitter **EDUCATION IN THE MICROCOMPUTER REVOLUTION**
Stewart Fraser **EDUCATION IN CHINA**
Jerrold Coombs **MORAL EDUCATION IN SECONDARY SCHOOLS**
Douglas Ray **ASSESSMENT OF MULTICULTURALISM IN THE MAINSTREAM**
Marilyn Westbury **RATIONAL MODELS AND TEACHER PREPARATION**
Shirley Grundy **TEACHERS' WORK: MAKING OR DOING**
John Laska **TEACHING AND CURRICULUM**
John Knight & Richard Smith **MULTICULTURAL POLICY AND A POPULIST RESPONSE**
Joseph Zajda **RECENT CURRICULUM REFORMS IN SOVIET SCHOOLS**

NEW EDUCATION

ORDER FORM

☐ **Please enter our subscription to *New Education***

We enclose $

☐ **Please enter my subscription under the terms of the Personal Subscription Plan** ***(please be sure to give your private address)***

We enclose US$

Please charge ☐ **MasterCard** ☐ **American Express** ☐ **Eurocard** ☐ **Access** ☐ **Visa**

☐ **Please send an inspection copy**

Name

Address

.....

.....

Signed Date

INTERCHANGE

A QUARTERLY REVIEW OF EDUCATION

- a leading international journal on the critical issues facing educators in Canada and abroad
- a recipient of the prestigious CAFE award for the most original work in educational foundations
- four issues per year feature scholarly articles, critical notices, book reviews, and debates
- recent special issues include Creativity, Thought, and Mathematical Proof; History, Science and Science Teaching; and Understanding Literacy (guest edited by David Olson)
- Edited by Ian Winchester

Regular One-Year Rates for 1990: Institutions $50.00; Individuals $35.00. For 1991: Institutions $55.00; Individuals $35.00. To order send cheque, purchase order, or complete MasterCard or VISA information to OISE Publishing, 252 Bloor St. W., Toronto, Ontario, Canada M5S 1V5.



Published by The Ontario Institute for Studies in Education, Toronto, Canada

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format

- Manuscript in triplicate.
- Typewritten or printed, double-spaced throughout.
- Abstract of approximately 100 words on a separate sheet.
- Manuscript not to exceed 6,500 words excluding graphics.
- Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.

Referencing

- Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s).
- Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format.
- All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
- Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes*; any citations in notes follow the same format as other references.

Graphics

- Each table and figure on a separate page at the end of the manuscript.
- Figures, charts, and diagrams in camera-ready form.
- Number tables in Arabic numerals and center over each heading, e.g., *Table 1.*
- Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1: ...*
- Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*

Style

- *AJER*'s editorial style conforms closely to the *Publication Manual of the American Psychological Association.*
- Editorial changes may be made to manuscripts.
- For spelling, consult *Webster's New Collegiate Dictionary;* spelling in quoted material remains as in the original.

Diskettes

If manuscript is accepted for publication and has been prepared using a Macintosh, IBM, or IBM compatible microcomputer, please send diskette. Microsoft Word, MultiMate, WordPerfect, WordStar, Xerox Writer, XyWrite, and DisplayWrite 3 **only** are acceptable.
Important: Please indicate program used.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index*; appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*



THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXVII NUMBER 2 JUNE 1991

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Thomas O. Maguire

Consulting Editors

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

Janet Donald
McGill University

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Lilian Katz
University of Illinois

Les McLean
Ontario Institute for Studies in Education

John McLeod
University of Saskatchewan

Linda Phillips
Memorial University of Newfoundland

Lauran Sandals
University of Calgary

Keith Sullivan
Dalhousie University

Phil Winne
Simon Fraser University

Ruth Wright
University of Ottawa

Faculty Advisory Committee

Harvey Zingle (Chair)
Fern Snart
Erwin Miklos
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Second Class Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXVII, Number 2, June 1991

EDITORIAL	ii	A Thought on Research Imperialism
KENNETH G. JACKNICKE WALTER D. SAMIRODEN	99	Some Perceptions of an Internship Program
BONNIE L. SHAPIRO	119	A Collaborative Approach to Help Novice Science Teachers Reflect on Changes in Their Construction of the Role of Science Teacher
DAHLIA BECK KERRY BLACK	133	Redefining Research Relationships: Two Heads Are Better Than One
CHUCK CHAMBERLIN JAN VALLANCE	141	Reflections on a Collaborative School-Based Teacher Education Project
COLLA JEAN MacDONALD	157	Elementary School Teachers Explain Why They Do Not Use Creative Dance in Their Classrooms
GLENDA H. CAMPBELL-EVANS	167	Nature and Influence of Values in Principal Decision Making
W. TODD ROGERS	179	Educational Assessment in Canada: Evolution or Extinction?

ISSN 0002-4805

EDITORIAL

A Thought on Research Imperialism

The Alberta Journal of Educational Research has served the Canadian educational research community for over 35 years. Generally speaking, we have published articles that have a national or provincial flavor with the goal of trying to encourage a distinctive research tradition. In the past few months I have had two occasions to reflect on this aim. Both involved seminars presented by prominent American research methodologists brought to our campus at great expense. In both cases the size of the audience was impressive, and it pleased me to see that so many people were interested in the study of educational phenomena.

Although both speakers came with outstanding credentials, neither knew much about the Canadian educational research community. Their reading apparently did not extend beyond their own journals, and their hosts had not informed them about the issues that currently burn in this country. Were I more cynical, I would have described it as an exercise in educational imperialism.

After the ritual apology about not being familiar with "what you people do up here," they spoke about research issues that currently dominate the agenda in the United States. The audience appeared to accept the transfer as though it automatically applied to the Canadian setting. Two examples come to mind. In the first case, the speaker spoke about how difficult is is to get qualitative educational research funded by national (American) funding agencies. She portrayed the qualitative research community as an epistemological underdog. This is demonstrably not the case in Canada. The Social Sciences and Humanities Research Council has provided generous support to several qualitatively oriented projects like the study of teacher thinking. The Canadian Society for the Study of Education has long had a strong group of qualitative researchers. Yet the impression left was that this is a general problem and that we would be well advised to devote some of our energy to correcting it.

The second speaker talked about the tyranny of testing programs and how the accountability movement warps the directions of true curriculum development. Although he may not have known it, the superficial similarities between the two federal systems are deceptive. Because the federal influence in education is much less in this country, because the traditions of educational administration at the provincial (state) level are so different, and because of the absence of a significant commercial testing industry in this country, the relationship between testing and curriculum is much different. Yet I feared that the audience left the talk with a feeling that the message applied to them.

Although the number of educational researchers in this country is small, there are many important thinkers. Their work has much to contribute to the solutions of problems that affect our educational systems. Isn't it time that we opted for contextual validity and sought advice from people who actually know something about the issues?

T.O.M.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 2, June 1991, 99-118*

KENNETH G. JACKNICKE

and

WALTER D. SAMIRODEN

University of Alberta

Some Perceptions of an Internship Program

This article summarizes an exploratory study of five beginning teachers' perceptions of a one-year internship. It addresses two themes revealed during analyses of interview transcripts: interrelationships among participants, and internship in relation to preservice teacher education. The interns' identities and status in the schools were mainly based on how they identified themselves and how they were identified through their association with the supervising teachers. Thus the relationship established between the intern and supervising teacher became the main determiner of whether the experiences could promote participant professional growth. As well, the interns believed that an internship program should be integrated with their preservice teacher education program.

Reality shock is the term used by Veenman (1984) to describe the experiences of beginning teachers as they make the transition from student status to full professional responsibility. Recognition of this period as being of key importance in the development of a teacher has led to the introduction of programs intended to provide support as the beginning teacher assumes the autonomy and authority of a professional. The concept of a transitional period of internship at the beginning of a teaching career is not new; Conant advocated the practice as early as 1963. But, following the reports of the Task Force on Teaching as a Profession (Carnegie Forum, 1986), the United States National Commission for Excellence in Teacher Education (1985), and the Holmes Group (1986) of deans of education which all endorsed the concept, an increasing number of programs are being implemented (Huling-Austin, 1987).

The literature on the subject of teacher internship reveals three major characteristics. First, a great deal of emphasis is directed at integrating theory and practice (Blackmore, 1968; Carney & Titley, 1981; Titley, 1984). Second, they are designed to "facilitate the transition from academic preparation to full professional responsibility" (Carney & Titley, 1981, p. 11). Third, notwithstanding these common characteristics in principle, there has been, and there continues to be, a wide range of programs all carrying the

Kenneth Jacknicke is a professor and Chair of the Department of Secondary Education. His interests lie in the areas of teacher education, curriculum theory, and science education.

Walter Samiroden is a professor in the Department of Secondary Education. His interests lie in the areas of teacher education, assessment, and science education.

label "internship." The structure of these programs shows considerable variation in duration, supervision, sponsorship, and the relationship to a teacher education program (i.e., pre- or postdegree).

Carney and Titley (1981) take issue with the labeling of a number of those programs as internships, preferring to call them extended practica. They note that these two terms are sometimes used interchangeably resulting in confusion. Their assessment of United States teacher internship programs led them to note that internship "follows other theoretical and practical aspects of preparation"; and, "is normally a terminal experience—the last stage before the granting of full professional status" (p. 16). This may indeed be the typical case, but many internships do not function in this way. Considerable latitude has to be granted in assessing whether these programs are intern programs.

Internships ranging in length from 16 weeks (Carney & Titley, 1981; Lang, Cornish, & Trew, 1980) to a complete school year may occur after completion of undergraduate studies (Slentz, 1978) or integrated within an extended undergraduate program (Jones & Barnes, 1984). During these periods, the interns may be supervised by university staff, school staff, school district consultants, or representatives of all of these, depending on the sponsorship of the program. Some programs make provision for the interns to return to the sponsoring institutions on a regular basis for discussion while others include on-site, inservice sessions (Allen, 1986; Harker, 1978; Mickelson, 1980; Silvernail & Costello, 1983). Some internships are prerequisite to graduation or certification, while others replace the probationary or induction year (Lawrence, 1985). Interns may be contracted and salaried by the local school board.

Most of the literature reveals investigations of programs that were incorporated into a university or college program leading to teacher certification, what Carney and Titley (1981) would probably consider to be a form of extended practicum. These studies usually examined some aspect of the field experience in relation to teacher development (McIntyre, 1983). It was not often, however, that much information was provided "about the structure of the student teaching experience itself beyond descriptions of when it took place, its length, and the number of classroom placements" (Zeichner, 1987, p. 101).

There have been some attempts, however, to investigate internships following completion of a degree. Silvernail and Costello (1983) compared traditional student teaching programs and internship programs for their impact on preservice teachers' pupil control perspectives, anxiety levels, and teaching concerns. Tabachnick (1980) described changes in graduate program interns who spent much of the time working in low-income area schools. Lawrence (1985) investigated peer evaluation in a one-year internship following graduation, and Slentz (1978) reported on a one-year graduate internship following teacher certification that, when combined with further study, could lead to an MEd degree.

One issue that needs further investigation regarding field experience programs is just how the experience itself is perceived by the participants. Zeichner (1980), in a paper on field experience in a preservice teacher educa-

tion program, argues that "the characteristics of field-based programs are not to be found in public statements of intention, but through an examination of the experiences themselves" (p. 53). When the opportunity arose to investigate a year-long internship program following certification, it was this notion of the experience of the participants that was considered by the researchers to be of paramount importance.

This article attempts to answer the questions: What interrelationships exist between the various participants of an internship program? and, What relationship exists between a postdegree internship and a preservice teacher education program? by describing an exploratory study of five beginning teachers' perceptions of their experiences in an internship. From the perspective of an intern, such relationships and established meanings are not static but are subject to constant readjustment as each individual strives to make sense of the world in which he or she works.

Conceptualizing Internship

The process of becoming a teacher has frequently been conceived as one of becoming more proficient at putting theory into practice. It begins with preservice programs in which theory is learned first and then practiced during student teaching. Fragmentation of teacher education into discrete program units has served to accentuate the belief that theoretical knowledge comes first, to be superseded by practical knowledge resulting from experience in the school context (Lanier & Little, 1986). Internship is seen as providing an extended opportunity for translating theory into practice, serving as a bridge between the two (Titley, 1984). The successful crossing of this bridge represents the completion of the transition from student of teaching to teacher. In this framework, the valued practical knowledge confers on the experienced teacher the status of expert in relation to the beginning teacher. The expert is able to provide support in the form of information, organization, demonstration, and facilitation. The supervisory nature of the expert-novice relationship ensures that any transmission of support is predominantly one-way.

An alternative framework to this linear progression from the theoretical to the practical rests on the assumption that theory and practice inform each other in a reciprocal mode during the process of becoming a teacher. There is a perpetual interaction among explanations, actions, and reflective experience. Regardless of context, teacher education is continuous and never complete: There is no conclusion to the phase of being a student of teaching. Internship is regarded in this model as the context in which the beginning teacher recognizes the interchange between theoretical and practical knowledge. Collegial relationships with experienced teachers provide the opportunity for reflection on the teaching act in a two-way exchange of critical analysis.

Too often beginning teachers in a traditional model come to believe that the way to learn more about teaching is through trial and error rather than through careful thought and scholarship. Lanier and Little (1986) are critical of this approach when they write "A model of field experience consistent with the liberal-professional approach to teacher education would strive to pro-

duce a deeper understanding of the way theoretical concepts from psychology, curriculum, and sociology are played out in classrooms" (p. 551).

They further support the notion of a liberal-professional approach, or an emphasis on the relationship between theoretical and practical knowledge, by indicating that as educators we should be providing "learning opportunities that will help prospective and practicing teachers acquire needed technique in ways that keep management in the background and student learning in the foreground" (p. 552).

Initiation to Teaching Project

In 1985, the Department of Education in Alberta announced the Initiation to Teaching Project or internship program. This project was to provide for the employment of up to 900 interns each year for a two-year period beginning in September 1985. Participating interns would not be employed as teachers, but would work under the guidance and supervision of experienced teachers selected by the principal of the school. The experience was to enable beginning teachers to refine their teaching competencies.

In order to be eligible, applicants must have resided in Alberta for at least three years, graduated from an Alberta teacher education program during the previous 18 months, and either possess or have recently applied for an Alberta Interim Professional Teaching Certificate. Further, they should have been unemployed or underemployed for at least four months and not have had previous employment as a full-time teacher.

Funding for the project was provided from three sources: Alberta Education (35%), Alberta Manpower (50%), and local school boards (15%) so that the intern employed through this program received a salary and benefits package of up to $15,600 for the school year. Alberta Manpower provided for a further grant of up to 75% of instructional costs to a maximum of $1,000 per intern for inservice and professional development.

Two principles were followed to guide the operation of the project.

1. The placement of an intern shall not result in the displacement of full, part-time or substitute staff.
2. The placement of interns and the program designed for each is a local responsibility. However, the program should have the following characteristics:
 a. guidance and supervision by certified teachers, selected by the principal, subject to approval by the superintendent (little if any professional preparation was provided for the supervising teachers);
 b. a planned program for the assumption of progressively greater responsibilities in the classroom (locally developed sequence);
 c. a full range of teaching experiences relative to the interests and abilities of individual interns;
 d. opportunities for participation in professional activities outside the classroom;
 e. clear identification of the expectations regarding the performance of the intern;
 f. provision of opportunities for the evaluation of the performance of the intern, consultation and remediation as required during the period of

participation (evaluation done primarily by the supervising teacher in consultation with the principal).

Information about the internship program was provided to prospective interns by Alberta Education and the Alberta Teachers' Association. Another source of information was through university classes and professors who sought information about the project. With regard to the individuals in this study, all had become interested in the project because they believed it was an excellent way to gain valuable experience without taking on full teaching responsibility. They also believed that such employment might lead to more opportunity for a full-time teaching position in the following year even though permanent teaching positions were not easily obtained.

Method

Werner and Rothe (1980) classified the methodology used for collecting and analyzing data in situational-interpretive studies such as this one into three phases: preentry, entry, and follow-up procedures.

Preentry concerns itself with the initial discussions with the participants. Because this type of research is based on conversation, participants' acceptance of the researcher is critical and entry must be planned carefully. The participants must have a clear understanding of the research and data gathering techniques to be used.

Following the decision by the two principal researchers to engage in the study, it was decided to limit the study to all of the secondary science interns in one major school system. It was restricted to one school board in order to limit variations in workshops available to the participants. The secondary science subject area was chosen because of the personal interest and academic backgrounds of the principal researchers, and because they: (a) knew the interns on a personal basis; and, (b) had a good working relationship with many science teachers teaching in the system. The possible sources of data included five science teacher education graduates who had completed their program during the previous year, five supervising teachers who had been identified by their principals, and five principals from the schools in which the interns were employed.

Permission was sought and received from the school board to carry out the study and to proceed with contacting the interns, their supervising teachers, and the principals of the schools involved. After providing a brief description of the program, an initial inquiry was made as to their willingness to participate. All the interns, supervising teachers, and four of the five principals agreed to participate in the study.

Following the initial contact by telephone, each supervising teacher and principal was visited by one of the researchers who outlined the investigation. Anonymity was guaranteed as much as possible, and participants retained the right to withdraw from the study at any time. Interview schedules were established that allowed for each intern to be interviewed four times: late October to early November, mid-December, mid-February, and early June.

The second phase, entry, involves data collection, and the method used was individual semistructured interviews. Prior to each set of interviews, the principal researchers developed the questions they wished to ask the par-

ticipants as guidelines for the interviews (see Appendix). The questions provided an outline of the topics for discussion but not necessarily the sequence. Freedom of movement within the interview situation was deemed essential to the process of data gathering, as was providing the opportunity for open-ended discussion on topics initiated by the participants. All interviews, averaging approximately one hour in length, were conducted at the schools and were tape-recorded.

Following each set of interviews, the researchers compared notes and prepared another set of questions for the next set of interviews using an initial interpretation of the previous interviews as a guideline. These preliminary interview analyses were validated with the participants by comparing the researcher interpretations with those of the participants at the beginning of the following interview.

Follow-up procedures involved transcribing all the tapes and conducting an analysis of the transcriptions. The tapes were transcribed by a research assistant and after several meetings of the researchers and the research assistant, it was decided to transcribe all the first interviews verbatim but to do some initial editing of the tapes while transcribing subsequent interviews (all tapes were saved for confirming details). The editing involved mainly deleting from the transcriptions introductory statements by the interviewers or conversation unrelated to the study, fillers such as, um, ah, well, and the utilization of some shorthand for common words and phrases. Emotional responses depicted by pauses or exclamations were retained in the transcripts by coding because of their significance in relating to the atmosphere, environment, and possible stress being felt by the participants.

During the initial interpretations, the researchers drew four major themes from the data: concerns relating to the administration of the program (e.g., funding, evaluation, salaries); teacher professional development (e.g., theory/practice relationships, teacher growth); preservice teacher education (e.g., relationship of internship to the BEd program); interrelationships among participants (e.g., expert/novice, hierarchial, collegial). The transcriptions were reread and coded according to the four major themes. Each theme was then analyzed to draw out subthemes and second-level interpretations. These interpretations were used as a means of coming to understand more deeply the phenomenon of internship as experienced by the participants. This paper addresses in part two of the four themes: interrelationships among participants, and preservice teacher education and its relationship to an internship program.

The Participants

All supervising teachers had been actively involved in preservice teacher education as cooperating teachers during practica and had at least 15 years teaching experience. They were directly involved in the intern selection process by interviewing candidates for the position at their schools. The principals were quite supportive of the intern project—two of them were successful in acquiring two interns for their schools. The principals had requested their staff members to present proposals for being supervising teachers and, in fact, encouraged specific members of their staff to submit

such proposals. Typically, they sought staff members who had been involved in the student teacher practicum programs. Once the supervising teachers were identified, they were directed to select possible intern candidates from a roster of applicants and interview them. The principals preferred the supervising teachers, who would be working most directly with the interns, to be the interviewers of the applicants and make the selection.

The intern participants in this study were five science education graduates ranging in age from 22 to 30 years, who had completed their Bachelor of Education programs during the previous year. Three interns were female. One of the women was employed at a senior high school (grades 10-12) while the other four interns were at junior high schools (grades 7-9).

The intern placed in a senior high school, Ann, was directly supervised by one teacher, but Ann also worked closely with four other teachers and taught at three grade levels. She was also heavily involved in extracurricular activities such as a drama club and as a coach for one of the girls volleyball teams.

Cal was at a junior high school where he worked with a teacher who had served as his cooperating teacher during one of his student teaching experiences. Cal had been requested to apply to do an internship and knew where he would be placed should he be interested. Cal and his supervising teacher had a mutual respect for one another from the beginning of the internship.

Betty was convinced that she wanted to do an internship immediately on learning of its availability. She had already experienced the stress of total immersion in a full-time position immediately after completing her training in a previous career and did not like it. She responded to a request to be interviewed by a prospective supervising teacher and succeeded in getting the position at a junior high school.

Edward had neither a job nor an internship at the onset of the 1985-1986 school year and had planned on being a substitute teacher. However, when an internship position became available during the middle of September, Edward applied for it and was accepted. The position that Edward filled had been vacated by an intern who had accepted a full-time teaching position in rural Alberta.

Doris had graduated from Education during the previous year and had completed six months as a supply teacher, experiencing success during a two-month assignment in one school. She thought the internship would provide her with a different experience than supply teaching and also a sense of permanence.

The Interns' Perceptions

Of Relationship to Self as Intern

In attempting to define their position during the first weeks of internship, it appeared to be easiest to simply relate it to some other profession or trade with which either they, or the people with whom they were interacting, were familiar. The availability of a common label, such as intern, was very convenient.

> To me, it's the same as an M.D. getting their doctorate ... he still has to intern. (Edward-1)[1]

> I've called myself an apprentice, because whenever I've said to people, I'm an intern, they say, what's that? (Cal-1)
>
> Even when I'm introduced [in another class], I was [supervising teacher's] intern. (Betty-1)

The acceptance of a student role seemed quite appropriate for the interns. After all, they had been students for most of their lives and had only recently graduated. On accepting their positions at their schools, all of the interns saw themselves as learners, or their supervising teachers put them into learners' roles, by controlling their involvement. The following excerpts illustrate this point:

> You get to watch a class in action with someone who is experienced, you get ideas ... you look at the ways one would handle a demonstration like an experiment. (Edward-1)
>
> Right now I take all the science classes, am making up units, teaching units—I still plan with [supervising teacher]—she tells me different approaches to try but allows me to put in my own approaches. (Betty-1)
>
> But, he's someone to talk to, to analyze, to give suggestions, to create scenarios that he obviously knows well or has gone through [experienced]—there's a lot of stuff to learn real quick and it [having a supervising teacher] makes the transition much easier. (Edward-1)
>
> I have the opportunity to learn from not only [supervising teacher's] ideas but also from my own. (Betty-2)
>
> When we started out, there's not that much responsibility and you know, I felt very much like a student. (Betty-3)

Seeing themselves as student teachers seemed to be natural. Some of these interns had just completed a round of student teaching a short four months prior to becoming interns. In Cal's case, he had returned to the school in which he was student teaching and his current supervising teacher had also been the teacher with whom he worked during student teaching. Quite simply, initial perceptions of their roles were those of student teachers.

> I see it more as a whole year of student teaching—it depends on who I'm working with—some students or teachers treat me as a qualified teacher—others treat me as a student teacher. (Ann-1)
>
> Didn't know what was expected of me—never could separate it from being a student teacher—found that I approached [supervising teacher] in the way I approached all of my cooperating teachers. (Betty-1)
>
> As a student teacher, you felt you had to go along, to a certain extent, and in the classroom, when you're watching someone's teaching style all the time, you tend to take that on—that's hard to separate here too and that bothers me because you want to be assertive and have your own teaching styles—but you tend to pick up on the style of the supervising teacher as I have. (Betty-1)

Of Relationship with Supervising Teacher

The interns were obviously influenced extensively by their interpretations of how they were viewed by their supervising teachers. After all, these were the

people who had contacted them, interviewed them and recommended that they be hired. The supervising teachers were also those who were given the responsibility to determine their roles; extend or delimit their experiences; provide a model which must be followed or encourage them to develop their own style; allow them to grow in stature, or maintain their subordination; and, finally, these teachers would evaluate their progress.

During the interviews with the interns, it became clear that they were anxious about the personal relationship that they might develop with their supervising teacher and that it was this personal relationship that would greatly influence the level of satisfaction that they would derive from their internship.

> It's a very personal sort of relationship between teacher and intern—it becomes that personal relationship that determines what the internship will be like. (Betty-2)

As the year progressed, it became obvious that the interns were experiencing varying degrees of independence. The supervising teachers could either hold power or relinquish it. In Doris' case, for example, it appeared that even though she was given some status as a certificated teacher at the beginning of the year, her supervising teacher controlled her level of independence during her internship and the level of independence or growth did not seem to increase as the year progressed. At mid-year, she appeared to be having some communication problems with her supervising teacher to the point where she was having difficulty in reading their relationship. By the end of the year, Doris felt that she had been manipulated.

> I was introduced as a part-time science teacher—just another teacher on staff. (Doris-1)

> [Supervising teacher] has given me a lot of independence—she listens from prep room but I can do what I want—she's here to guide. (Doris-1)

> Christmas will be good for [supervising teacher]! I think [supervising teacher] needs a break! So I don't really know what's going on with ——, it seems to me that —— is under a lot of pressure right now. ...Yesterday we had a good little talk, we laughed—it's been quite a while since we laughed together. (Doris-2)

> I was manipulated and I suppose toward the end of the year, I should just "have it out" with [supervising teacher] and "clear the air." Probably controlled in a sense. (Doris-4)

It appeared that Betty's status was, if anything, even more controlled than Doris' by her supervising teacher. She started her internship with enthusiasm but did not seem to gain status on her own or from the supervising teacher. She remained as a student, by and large using her supervising teacher's materials. She was controlled by the person who determined her activities.

> As far as incorporating anything into the unit, I feel I'm free to do that—my own ideas—I do use some of [supervising teacher's] material right now, worksheets etc. I don't feel like going in there and changing everything because of my position as an intern. (Betty-1)

> I'm beginning to feel more like a teacher but yet at the same time I feel like a student because I can never really assert myself toward being a teacher. (Betty-4)

> I'm sitting there and [supervising teacher] won't [let me substitute]—she gets very upset if I watch someone else's class for maybe a period—that's [no subbing] her expectations—maybe I lost something there because I was under one person. (Betty-4)

Ann's internship experiences began in a different way from all the others and it had its stormy times. She started her year working directly with one supervising teacher and indirectly with four others preparing for five different classes. This arrangement was through mutual agreement and, to a large extent, was a result of her desire to work with a number of different people so that she could make her experience as broad as possible.

> [Supervising teacher] asked me "What would you want to do?" and I said I wanted to work with a number of different teachers and areas—even areas I don't have a background in, to see different teaching styles and interrelations between students and teachers—so it was decided. (Ann-1)

Her relationship with her main supervising teacher was rather loose, and Ann seemed to have the freedom to accept or reject requests from other teachers for her assistance. She soon found herself overburdened to the extent that she wanted to set up a meeting with all her supervising teachers. It was quite obvious that her supervising teachers had either allowed her to gain the status and confidence necessary to make such a request, or through their interaction Ann felt that she would not be viewed negatively by taking the initiative in calling such a meeting.

> We set up a meeting—I sent around "Happygrams" inviting all the teachers I work with ... so everybody would know what my position is and what my commitments are. So they understand why I have to say "no" sometimes.... the meeting clarified things.... a lot of them said I was doing too much—more than a first year teacher. (Ann-2)

In reflecting on the year's activity, Ann saw herself as being very fortunate in having so broad an experience.

> I think it would be easiest to have just one person to work with—I would know exactly what he was like and develop a good working relationship. Whereas I think I'm much luckier having worked with five teachers—each of them have different styles and personalities—not only getting to know them personally but also their teaching styles.... The people I'm working with have excellent questioning techniques, good wait-times, very good people—I guess I'm working with the "cream of the crop." (Ann-3)

Edward and his supervising teacher became known as a team in their school early in the year. This was not surprising because there already were teamed teachers in the school in mathematics and social studies. Another team was not seen as a novelty. Being labeled as a team seemed to encourage them to perform as a team. Together they planned ways in which they would either accept or reject requests for Edward's time or assistance. Whenever either was asked for Edward's help, that person always had to check with the

other's planned activities. Substitute teaching was occasionally accepted if the situation seemed to warrant acceptance and the experience was seen as a broadening one. They discussed teaching learning situations as a team and seemed to develop a mutual respect.

> They [students] tried to play [supervising teacher] and I against one another, and we find that kind of enjoyable. (Edward-1)

> We can laugh when we "blow it"—a working friendship—pretty close. (Edward-3)

> I'm very comfortable with the man, I'll probably keep a friendship with [supervising teacher] the rest of my life whether I stay here or not. I've got a lot of respect for him as a professional—his interest lies with students. (Edward-4)

Cal's situation was unique as far as the different internships were concerned: he was with a former cooperating teacher from student teaching days. They knew one another and seemed to start their relationship on a more collegial level.

> We knew each other before and we had some experiences with each other before, I'm enjoying a great amount of freedom and confidence that [supervising teacher] has in me. (Cal-1)

They established flexible guidelines through mutual agreement and were directed by those guidelines.

> I like the idea of a mentor, he's the person who I can approach at a moment's notice; he's the person to help me with professional growth. (Cal-1)

The supervising teacher, however, maintained a superior status by being the final decision maker with regard to his intern's involvements in other school activities or responsibilities. Cal saw this as being a way in which he was protected from getting overburdened.

> [Supervising teacher] has made it explicit that if there are any changes to be made to my program they must be through him and I second that fully. (Cal-2)

Of Relationship with Students and Classrooms

The interns were located in schools where the students were already experienced in having student teachers. The interns realized that in the eyes of the students they behaved and appeared as student teachers. They observed the cooperating teacher's classes and even sat at the desk where the student teachers normally sat.

> At the beginning, they didn't know who I was—a student teacher—perhaps because I took the student teacher's desk. (Doris-1)

The students, already familiar with the four-week periods of student teacher practicum programs, recognized that the interns were something different when they were still at the school after four weeks had passed.

> At first they were very confused ... they were kind of used to people being in the classroom because of student teachers. But they found out—"Hey, four

> weeks are up, the guy hasn't left us!"—they finally got the idea and they kind of like it. (Edward-1)
>
> They [students] didn't know who I was and I think now they do—I'm a teacher here. So I've kind of risen above what I was, at the beginning. Also how I see myself too, because sometimes they'd ask me "Are you going to be teaching here next year?" So, they see me as a teacher. (Ann-4)
>
> Also, the students have a better understanding of who I am. It's nice to have a student say I'm their favourite teacher and I'm the intern here—that's a real feather in my cap! (Ann-4)
>
> I've got a homeroom and at first they were confused but *now* I'm a teacher on staff to them—that is sort of part-time. (Doris-1)

For the participant interns, recognition of their status as above that of student teachers was important. It signified a separation of the stage of preparation from the beginning of practice, or a movement from student teacher to teacher. Accompanying this perceived rise in status, some participants experienced a sense of authority and power:

> So right now [after two months], when it comes to assignments, when it comes to marking, I'm the final authority. I'm the teacher, and I'm perceived as such. (Cal-1)
>
> For the most part that was my class, and now when we're doing marking [the supervising teacher] said I could have done it on my own, because the [supervising teacher's] sort of out of touch with it and they've been coming to me. (Ann-3)
>
> On the whole, the staff will actually consult me about [the classes I teach]. (Cal-4)
>
> Relationships [with students] develop where you could use them as "controllers" in the classroom and he did all the work for me as far as behavior was concerned. (Edward-4)
>
> We'd both be in the classroom and often times we'd both speak at different times, and we were very good at that, co-instruction really went well. But in terms of "whose classroom is this?"—I had to straighten that matter out, and to really get into the teaching act, I had to own the classroom. It had to be my family. It had to involve individuals who had to answer to me and to whom I had to answer. (Cal-4)
>
> With the administration, everything goes through [my supervising teacher]. Yesterday, I had a parent call for me—boy, did I feel good—I'm not a teacher, I'm someone with [supervising teacher] who is the teacher. (Betty-4)

Internship and Preservice Teacher Education

A major theme that seemed to grow out of an analysis of the transcripts focused on the relationships of theory and practice. Much of the writing in recent years regarding educational theory does not make direct reference to practice. It is this dichotomizing of theory and practice that has led to the "mode of intellectualism that is so dominant in our culture, such that only a few see difficulty with [the statement] one should know theory first and then apply it to practice" (Aoki, 1984, p. 12). However, a growing number of

scholars on the North American scene are disillusioned with logical positivism and with the theory-practice dichotomy it represents.

It became evident that the interns were, by and large, steeped in a mode of thought that exemplified the dualism of theory and practice. Yet many of their comments seemed to suggest a "coming back" to their previous educational experience (preservice education) in order to better understand their current situation. It appeared that in real-life situations (i.e., internship), theory and practice were opposite sides of the same coin (praxis). Groome (1980) states it as:

> To understand praxis requires a shift in consciousness away from dichotomizing theory and practice, towards seeing them as twin moments of the same activity that are united dialectically. Instead of theory leading to practice, theory becomes or is seen as the reflective moment in praxis, and articulated theory arises from that praxis to yield further praxis. (p. 152)

It may be that internship might be approached as a praxiological act. In Freire's (1972) words, "Praxis is reflection (thought) and action (practice) upon the world in order to transform it." Freire is critical of the dualism created by separating theory and practice which can result in action without thought (anarchy) or thought without action (verbalism). It appeared from our reading of the transcripts that what the interns were actively engaging in was purposeful and reflective action by which knowing arose through engagement in a social situation.

Some educators suggest that theory becomes more meaningful after having some experience; it is quite common for student teachers to separate theory from practice during their teacher education program when they take methods courses along with their practica, unless the course assignments demand some integration. Edward saw a lack of practicality in university coursework, whereas other interns either recalled information from their coursework that became pertinent to their classroom situations, or had experiences emphasizing either the separation or the integration of theory and practice during their internships.

> I'd like to see a little more emphasis [in courses] on the mechanisms of teaching.... implementing ideas into the classroom. (Edward-1)

> Often I go back and look at my notes ... refer a lot to [curriculum and instruction] book with teaching skills in it. (Betty-1)

> One of the things C & I [curriculum and instruction] did teach me ... was the actual preparation; organization ... theories as well—at the time I thought it was a waste, but ... I bump into them or end up using them ... I wish I had spent more time discussing them. (Ann-1)

> Then when I came out, I started seeing connections and ties [with university courses]. It would have been more useful if I'd had a bit of experience and then some theory and then gone back out again. (Ann-4)

> There's no time to reflect in your CI's and then you're going out into your teaching rounds and coming back and it's just a little too rushed but generally in the overall picture, they're very good. I'm now getting the benefit of those by going back and looking at the materials that I had. (Edward-1)

Toward the end of their internship, when the interns were asked to consider how such an internship as they had experienced might relate to an optimal teacher education program, numerous aspects emerged.

The need for integrating internship experiences with curriculum and instruction courses was emphasized.

> You gotta have the C and I [curriculum and instruction courses] definitely, but maybe you could incorporate the CI in with an internship where they do go back and then there's times here [school] when you're definitely not needed. (Betty-1)

> A lot of theory—you don't *really* understand it until you've actually gone out and taught. Then when I came out I started seeing connections and ties. It would have been more useful if I'd had a bit of experience and then some theory and then gone back out again. (Ann-4)

For some, the internship was seen as being much more important than student teaching, whereas others saw student teaching as an important kind of prerequisite to internship, and that an internship would produce more proficient teachers than a full-time first-year teaching position.

> I'm really glad we had the opportunity to experience a practicum before the internship, because it's almost like a testing ground to see if you could handle the internship ... if you don't have that then you may as well not continue. (Ann-1)

> I think the internship program, philosophically and practically, is a logical prerequisite to becoming a proficient first year teacher. (Cal-2)

> I've gotten a more rounded experience ... I've worked with more people. I suppose a first-year teacher would do that to a certain extent, but I think I have less tunnel vision. (Ann-3)

When considering how or where an internship might fit into a teacher education program, the interns were not in total agreement. Doris believed that it should be separate from the university program, whereas the other interns preferred to see it associated with their university developed teacher education program. During discussions of this topic, the relationship of internship to various other aspects of becoming a fully certificated teacher surfaced. Questions that became important included: How should the time devoted to an internship be recognized? Should it be seen as pensionable service? Should the internship be a requirement for permanent certification or should it be optional? and, How long should an internship be?

> I'm not really gung-ho about making it a year on your pension and increments, etc. I'd much rather see it attached to university and be a qualifying year that elevates your qualification as a teacher. (Cal-2)

> And it [internship] should possibly be the final year or most of the final year with the concluding section at the university. I wouldn't take away the student teaching. (Cal-2)

> I would love to see the internship before a person graduates and it would eradicate a lot of things like the pay. It's [the pay] more acceptable when you

> don't have a degree in your pocket. Any negative thoughts that I have are for that reason. (Betty-1)
>
> I think it would be great if it [internship] was part of teacher education. (Betty-1)
>
> Use the internship toward permanent certification—do it exactly as they have it now, out of university, have an interim teacher certificate, intern for a year with that document, which makes you legally responsible for the welfare of the kids, but you are still supervised and then you are granted a permanent teaching certificate—make it only a year. (Doris-4)
>
> I think it should be compulsory and I think that they could add a fifth year and make it a year of teaching and it should be paid and it should count towards permanent certification if it's a successful year. (Ann-4)
>
> What I would like to see it as a 6 or 8 month program within a year. Until you've done an internship, you get a temporary certificate but you cannot get your permanent status until you've completed that internship. (Edward-4)

Discussion

The Initiation For Teaching Project

Through our involvement in this research and in particular our discussions with the participants, it became clear that this project was hurriedly conceived and initiated after only a short gestation. Neither the interns nor the school personnel participants had clear understandings of their roles. For example, most of the interns wrongly perceived that their involvement in the project would lead to a better opportunity for gaining full-time employment as a teacher in the following year. The suggestion that the intern was to be evaluated promoted the development of a student teacher-supervising teacher relationship among some participants. This was totally incongruent with the fact that the interns already had been certificated and, as such, could have sought employment as fully qualified teachers. In retrospect, this initiation to teaching project can be described as a too rapidly implemented make-work project at a time when newly certificated teachers were not in high demand.

The Identity of an Intern

For the participating interns in this study, their identity as interns was problematic. In terms of the temporary nature of the position, the reduced teaching and administrative responsibilities, and the supervision by an experienced teacher, the similarities to their recent student teaching experience appeared overwhelming. As with student teaching, they both identified themselves and were identified by their association with the supervising teacher. They saw themselves as students needing to learn "real quick," to be "told" different approaches and to "go along with" the supervising teacher. Many of their concerns focused on how they fit into the teaching situation, paralleling stages described by Fuller and Bown (1975). The setting of the parameters for instruction by some supervising teachers maintained the student teacher dimension for their interns. For others, the classroom students and other teachers contributed to their recognizing themselves as teachers. Some social interactions with students and teachers may be per-

ceived as validating experiences (Jordell, 1987) in which the beginning teacher progresses toward an acceptance and affirmation of his or her situation as teacher.

The interns' perceptions of their individual situations revealed a desire for autonomy not built into an internship program that is expected to bridge an artificial gap between theory and practice. Autonomy requires understandings brought about by a mutual informing of theory and practice by the intern, and not interceded by the supervising teacher. For some participants in this study, the satisfaction of demonstrating responsibility through taking independent action shone through the haze of supervision.

The Nature of Support

In this study, the perceptions of the interns revealed predominantly hierarchial superior-subordinate relationships between themselves and experienced teachers. In becoming acquainted with the ways of acting in a particular school setting, there is an inherent one-way flow of information from those who already know to those who are learning. Internship may be regarded as a phase in which the novice is inducted into the teaching profession as a receiver of the rules of the game. For the interns, as with student teachers, this emphasis "often means a continuation of the teaching practices by which they were taught as well as the tendency to see the prevalent patterns of teaching as the only ones possible" (Lanier & Little, 1986, p. 551). Such modeling experiences by the interns are unlikely to make much contribution to their professional growth or those of the other of the participants, especially the supervising teachers.

The interns in this study noted the key role of the relationship with their supervising teachers. For one of the participants (Cal) the relationship was perceived as that between mentor and protégé, in which the mentor nurtures the social and intellectual development of a less experienced colleague, rather than the hierarchial relationship alluded to above. Galvez-Hjornevik (1986) suggests that the mentor-protégé pair connotes a voluntary and deep relationship "not limited to basic direction and encouragement" (p. 10) and as such, it is probably optimistic to expect such quality in teacher internship programs without some built-in program of preparation to help this come about. As Thies-Sprinthall (1986) has pointed out, experienced teachers are not necessarily effective mentors. She argues that there is a need for intensive preparation for the role so that the mentor is able to contribute an appropriate balance of support and challenge to the relationship. Unfortunately, in the program described in this study, there was no preparation for the supervising teachers in their roles as mentors. Among the outcomes of such a preparatory program for mentor teachers described by Thies-Sprinthall (1986) was a reassessment of their own classroom performance. The mentors all thought they were effective teachers prior to the program, and all decided after the program that their initial levels were often only partially effective (p. 18). Their ability to analyze and evaluate shifted from the lesser to greater complexity. In learning how to become mentors, these experienced teachers had grown in their understanding of the reciprocity of theory and practice.

All teachers, both beginning and experienced, are in the process of constructing and reconstructing meanings for their situations. Instead of becoming more proficient at explaining the teaching experience, the teacher becomes accustomed to seeing the experience as problematic. He or she then has the opportunity to engage in a reflective interchange with the situation, what Schön (1983) has called a "reflective conversation," in order to consider what possibilities exist for action. A collaborative inquiry into the overlapping situations of intern and mentor-teacher could be regarded as the intersection of reflective conversations. As colleagues, the two teachers attempt to make explicit the interplay of theoretical and practical knowledge in each situation. Through dialogue, the partners construct meanings for the problems at hand, rejecting any imposition of meanings through taken-for-granted routines or labels.

And so internship can be regarded as an opportunity for two-way support in a collegial relationship. Andrews (1987) has suggested a "confluence paradigm" for the linking of induction and inservice programs. Collegial support between teachers at varying stages of their development could become the basis for change. However, such exchanges must rest firmly on the duality of theory and practice. Too much emphasis on experience as simply practice, without theory-based questioning of the practice, may give rise to the sharing of introspections that actually inhibit the ability to see the range of possible decisions and actions (Buchmann & Schwille, 1983; Floden, Buchmann, & Schwille, 1984; Lanier & Little, 1986).

Recommendations

As a result of this study we believe that if the following recommendations were implemented, they would help develop a higher quality internship program.

1. That, prior to implementation, any internship program should be carefully conceptualized with input from representatives of all interest groups. The focus of the internship should be on the promotion and support of professional growth of all participants—in particular, the interns and the mentoring teachers.
2. That an internship program should be centered on the notion of collegial relationships among the participants, rather than on a notion that presents an expert/novice or hierarchical relationship. This would help to substantially reduce the separation of theory from practice and promote the development of the reflective practitioner. One way to promote the growth of collegial relationships would be to involve supervising teachers in an intensive preparation such as that suggested by Thies-Sprinthall (1986). The preparation would be designed to assist participating supervising teachers develop an understanding of mentoring rather than supervising.
3. That the internship program should be integrated with the interns' preservice teacher education program. Again, this would help the participants avoid the artificial separation of theory and practice and promote the development of more liberal-professional career teachers (Lanier & Little, 1986) where teacher education is considered to be con-

tinuous and lifelong. It is probable that such an integration would also help to incorporate more of teachers' practical knowledge into those university-based courses that are at present primarily theoretical.

Note

1. Each intern was interviewed on four occasions throughout the school year. The number indicates the sequence of interviews.

Acknowledgment

The researchers acknowledge the assistance of P. Rowell in the preparation of this paper.

References

Allen, S. (1986). Quality control: The key to the success or failure of an internship program. *Education Canada, 26*(3), 38-40.

Andrews, I. (1987). Induction programs: Staff development opportunities for beginning and experienced teachers. In M.F. Wideen & I. Andrews (Eds.), *Staff development for school improvement*. Philadelphia, PA: Falmer.

Aoki, T. (1984). Curriculum implementation as instrumental action and as situational praxis. In T. Aoki, T.R. Carson, B.J. Favaro, & L.M. Berman, Understanding situational meanings of curriculum inservice acts: Implementing, consulting, Inservicing. *Monograph No. 9* (pp. 3-17). Department of Secondary Education, University of Alberta.

Blackmore, D.S. (1968). Integration of theory and practice in internship. In *Internships in teacher education: 47th yearbook of the Association for Student Teaching* (pp. 41-63). Washington, DC: Association for Student Teaching.

Buchmann, M., & Schwille, J.R. (1983). Education: The overcoming of experience. *American Journal of Education, 92*, 30-51.

Carnegie Forum on Education and the Economy. (1986). *A nation prepared: Teachers for the 21st century*. Report of the Task Force on Teaching as a Profession. New York: Carnegie Corporation.

Carney, R.J., & Titley, E.B. (1981, July). *Internship*. Paper presented to the Internship Ad Hoc Committee of the Board of Teacher Education and Certification, Edmonton.

Conant, J.B. (1963). *The education of American teachers*. New York: McGraw-Hill.

Floden, R.E., Buchmann, M., & Schwille, J.R. (1984). *The case for separation of home and school*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Freire, P. (1972). *The pedagogy of the oppressed*. New York: Herder and Herder.

Fuller, F.F., & Bown, O.L. (1975). Becoming a teacher. In K. Ryan (Ed.), *Teacher education* (pp. 25-52). National Society for the Study of Education, 74th yearbook, Part II. Chicago: University of Chicago Press.

Galvez-Hjornevik, C. (1986). Mentoring among teachers: A review of the literature. *Journal of Teacher Education, 37*(1), 6-11.

Groome, T. (1980). *Christian religious education: Sharing our story and vision*. San Francisco, CA: Harper & Row.

Harker, W.J. (1978). The Saanich experiment in internship. *Teacher Education, 13*, 86-91.

Holmes Group. (1986). *Tomorrow's teachers*. East Lansing, MI: Author.

Huling-Austin, L. (1987). Teacher induction. In D.M. Brooks (Ed.), *Teacher induction—A new beginning*. Reston, VA: Association of Teacher Educators.

Jones, A.H., & Barnes, C.P. (1984). The California consortium: A case study on seeking change in teacher education. *Journal of Teacher Education, 35*(6), 5-10.

Jordell, K.O. (1987). Structural and personal influences in the socialization of beginning teachers. *Teaching and Teacher Education, 3*(3), 165-177.

Lang, H., Cornish, W., & Trew, L. (1980). Partners in training: How it works and why. *Education Canada, 20*(3), 14-21.

Lanier, J.E., & Little, J.W. (1986). Research on teacher education. In M.C. Wittrock (Ed.), *Handbook of research on teaching*. New York: Macmillan.

Lawrence, D. (1985). The Toledo plan for peer evaluation and assistance. *Education and Urban Society, 17*, 347- 354.

McIntyre, D.J. (1983). *Field experiences in teacher education*. Washington, DC: Foundation for Excellence in Teacher Education and the ERIC Clearing House on Teacher Education.

Mickelson, N.I. (1980). Practice teaching: A comment and a partial solution. *Canadian Journal of Education, 5*, 87-90.
National Commission for Excellence in Teacher Education (1985). *A call for change in teacher education*. Washington, DC: Author.
Schön, D.A. (1983). *The reflective practitioner*. New York: Basic Books.
Silvernail, D.L., & Costello, M.H. (1983). The impact of student teaching and internship programs on preservice teachers' pupil control perspectives, anxiety levels, and teaching concerns. *Journal of Teacher Education, 34*(4), 32-36.
Slentz, K. (1978). Graduate internship for teachers. *Teacher Education, 12*, 92-96.
Tabachnick, B.R. (1980). Intern-teacher roles: Illusion, disillusion, and reality. *Journal of Education, 162*(1), 122-137.
Thies-Sprinthall, L. (1986). A collaborative approach for mentor training: A working model. *Journal of Teacher Education, 37*(6), 13-20.
Titley, B. (1984). The concept of internship in teacher education. *Teacher Education, 24*, 84-93.
Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research, 54*(2), 143-178.
Werner, W., & Rothe, P. (1980). Curriculum praxis: Doing school ethnography. *Monograph No. 2*. Department of Secondary Education, University of Alberta.
Zeichner, K. (1980). Myths and realities: Field based experiences in preservice teacher education. *Journal of Teacher Education, 31*(6), 45-55.
Zeichner, K. (1987). The ecology of field experience: Toward an understanding of the role of field experiences in teacher development. In M. Haberman & J. Bacchus (Eds.), *Advances in teacher education* (vol. 3). Norwood, NJ: Ablex.

Appendix

Questions Aimed at Initiating Dialogue for Two Intern Interviews

Internship—Interview #1

1. How did you first get involved in becoming an intern?
2. How did you become aware of it (internship)?
3. What were major influencing agents that led you to apply?
4. What kinds of changes have taken place in your thinking about internship over the past seven or eight weeks?
5. How do you feel about your choice to become involved in the internship?
6. How would you compare the internship with a first year teaching position?
7. How were your responsibilities decided upon?
8. How do you feel the students perceive you?
9. How are you perceived by the other teachers in the school (other than your supervising teacher)?
10. How does the feedback you receive now compare with feedback you received when student teaching?
11. How would you compare your internship with student teaching?
12. You've been at it (internship) for two months now, what are some of the most important learnings?
13. Do you see the internship fitting into a teacher education program?

Internship—Interview #2

1. What is your relationship with the students like now?
2. In this last while, do you sense any changes in your relationship with the other teachers?
3. As an insider, do you sense any changes in the relationships between your supervising teacher and the other teachers?
4. After nearly three months, how would you describe an ideal internship program?

5. Would a master teacher be making decision as to what your experiences should be?
6. How do you think an internship compares with a first year teaching position?
7. Have you noticed any change in how you're perceived by other professionals around you?
8. On the notion of legal responsibilities—what do you perceive as yours?
9. How would you feel about your role as an intern now if you were told that your professional responsibilities were nil, regardless of whether or not you are a certified teacher?
10. How would you see an internship relating to your precertification program? Your university program?
11. What other kinds of things have been highlights since the last time we talked?

Note: The above questions only served to initiate dialogue on a variety of topics during the interviews. Other topics developed and provoked different directional questions during the dialogues with each participant in relation to their responses and apparent interests associated with their very individual situations. Each intern was interviewed on four occasions throughout the school year. The number indicates the sequence of interviews.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 2, June 1991, 119-132*

BONNIE L. SHAPIRO
University of Calgary

A Collaborative Approach to Help Novice Science Teachers Reflect on Changes in Their Construction of the Role of Science Teacher

This article describes an approach to the study of growth and change in student teachers' views about science teaching as they gain experience and practice in the field. Views about the thought and activities of science teaching were assessed prior to students' university teacher methods courses, then again following their classroom practicum teaching experiences. Student teachers provided linguistic categories (constructs) for various science teaching experiences, then collaborated with interviewers to convey their ideas about the reasons for changes that occurred in the application of these categories. The research approach allowed the collection of both objective (survey) and subjective (interview) data. Analysis of the types of personal construct changes that occurred allowed the delineation of categories of change as students became socialized in the role of teacher. Excerpts from individual case study reports illustrate the outcomes of the study and are the basis for considering the implications of incorporating student views in the design of teacher education programs.

> I don't know why there is so little effort to find out who the future teachers are before they come to the institution, or so little effort to socialize them into teacher education. (John I. Goodlad, in Watkins, 1990, p. A17)

Introduction

A new thoughtfulness is emerging in educational research. We are seeing the importance of directing our attention to knowledge as it develops within the person. In the field of teacher education there is concern with issues of individual coherence and integrity of understandings that are held by preservice teachers about teaching and learning. New research approaches are needed that allow insight into how teachers reflect and grow in their work. Goodlad's concern places a spotlight on the need for fundamental work in our knowledge of future teachers and the ways that we may guide their socialization into the profession.

In this article, an approach is described that has been developed to understand the changes that novice secondary science teachers experience during the first portion of a teacher education program. Several questions guide this ongoing research: What is the nature of the knowledge student teachers acquire in the process of learning to teach? How do student teachers change in their ways of thinking about teaching? What changes in growth and

Bonnie Shapiro is an associate professor jointly appointed by the Departments of Teacher Education and Supervision and Curriculum and Instruction. She has expertise in science and language education.

development are apparent to the student teachers themselves? These questions are of enormous theoretical and practical interest. The answers will help us to develop understanding of changes that occur for beginning teachers and how we might assist in their growth and development.

The problem of understanding student teacher knowledge acquisition is complex. Osborne and Gilbert (1985) note, "all teachers have views of learning which are implicit in their practice, but are rarely articulated, even to themselves." Such views may be even more difficult to access in the case of the student teacher. Student teachers are only beginning to clarify their views about science teaching and learning. They come to the student teaching experience with a wide range of beliefs and perceptions about what it means to be a teacher (Hewson & Hewson, 1988). Many changes occur in students' thinking during the student teaching practicum. This is usually a time when science teacher educators are not present to interact with student teachers. Teacher education programs are the first chance we have to begin to know our students, yet there is little opportunity to observe the ways that students make connections using the resources and ideas that teacher educators have attempted to make available to them. This difficulty may be largely due to traditional structuring of teacher education programs. Large classes impede knowing individual students well. We require new approaches to understanding changes in student teachers. Calderhead (1987) asserts that the processes of learning to teach are complex and not adequately conceptualized. He states further (1989) that "empirical research on student teachers' knowledge and thought processes and how these are influenced by alternative approaches and designs in teacher education might enable us to test out both the realities and possibilities of reflection in teacher education" (p. 44)

The Need for Collaboration in Reflective Practice

The central theoretical position of this work is that learning to teach is an active process requiring opportunities for interaction between thought and actions. The value of "reflection-in-action" (Schön, 1983) has been an important recent focus in the study of teaching practice (Grimmett & Erickson, 1988). It has been suggested that Schön attributes special meaning to this phrase (Grimmett, 1988), that is, the practice of reflection should occur in the action, and that this reflection is a kind of research oriented to practice. Schön (1988) writes:

> Research should be oriented to practice. Instructional supervision would be seen, on this view, as a research activity. Both the reflective teacher and the reflective coach are researchers in and on practice whose work depends on their collaboration with each other. It is the nurturing of groups of researchers of this sort, at the core of the schools of education, that I believe holds greatest promise for healing the breach between educational research and practice, revitalizing the schools of education, and mobilizing their resources in support of reflective teaching in schools. (p. 29)

Recent rethinking in the provision of teacher education practicum experiences has placed value on the personalized, reflective opportunities needed to enhance the student teacher's development and growth. Although teacher

educators attempt to provide experiences that allow students to reflect on and reorganize their thinking about teaching, institutional structures have not reflected the same pace of change. The resources needed to allow this to happen are not always available.

The Opportunities of the Student Teaching Experience

Fundamental to design of the research are several beliefs about educational practice. First, the process of becoming a teacher should be an occasion to value and clarify the principles and values that underlie one's own ideas about what it means to be a teacher. It should be a time of articulating one's images of teaching and of acquiring language for communicating ideas about teaching and learning. It should be a time for exploring ideas about interacting with students, about using curriculum resources, and about being in schools. It should offer a chance to explore many different methods of teaching. It should be a time of developing a deeper awareness of oneself in the task one is undertaking. Research into one's practice means valuing one's present ideas and teaching self. If this process is to be personally meaningful, it is from such opportunities that reorganization and reconceptualization of self as teacher can effectively take place.

Foundations of the Research in Personal Construct Theory

One might consider the experience of student teaching to be like the solving of a problem, or of a series of problems, at times even several problems at once.

In the view of personal construct theory, we straddle the unknown with what we know. If we are to consider the dilemmas of student teachers, it is valuable to consider what it is that they already know and what *they* consider actually to be a dilemma. This study explored what it is that students set for themselves as dilemmas at the beginning of their experience and what they noticed that changed for them, became new, or surprised them. We are interested in understanding the ways that students use what they know—how they see themselves using the material from the experiences we offer them to help straddle the unknown with what they know.

At the heart of the study is the belief that the chief participant in the experience, the novice teacher, should be deeply involved in the discussion that determines the meaning of the experience. The meanings attributed to the experience by the person actually having the experience, the novice teacher, are given high status. The project is rooted in a constructivist theoretical position concerning the development of teachers (Erickson, 1987; Shapiro, 1988). The principles of personal construct theory (Kelly, 1969, 1970) are used to explain the ways that student teachers understand changes in their experience. Personal construct theory provided a framework for helping the student teachers themselves analyze the meaning of changes in novice teacher thoughts about their own growth and development. The approach taken in the research permitted involvement of the student teacher in a systematic discussion about changes in their own development that they might not otherwise have been able to see.

Conduct of the Study

A reflective, collaborative interview was developed to elicit personal constructs and to allow interviewers and student teachers to come to know one another. Nine interviewers worked with 23 volunteer student teachers before, during, and following the first phase of their teacher education program. This consisted of seven weeks of methodology coursework followed by four weeks of practical classroom experience.

Survey information was collected and interviews were conducted prior to and following students' university based methods course (seven weeks) and their initial practical teaching experience (four weeks). Each novice teacher worked closely with an interviewer who assisted the student teacher in the development of sets of personal constructs. Personal constructs are essentially linguistic structures used by the individual to anticipate new situations, guiding behavior in a new or unfamiliar area. In this case, the new event was taking on the role of science teacher. Following the first seven weeks of coursework and the practical teaching experience, student teachers were guided to reconsider their thinking about the events of science teaching. Collaboration with a caring guide, the interviewer, allowed the review of the use of personal constructs, the novice's ways of categorizing the thoughts and actions of science teaching.

It is helpful to show the kind of distinction made by personal constructs. As an example, a student teacher considered "gathering laboratory equipment" to be an administrative task, whereas "conducting a demonstration" was considered to be a teaching task. The phrase *administrative task*, the student's term, is placed in opposition to *teaching task*. These two terms represent a personal construct. Once constructs are established, other events of teaching may be categorized using these terms.

The student teacher rated each task along a continuum linking the two constructs before and after the initial portion of the methodology course and the practicum experience. Significant changes in the student teacher's use of the original categorizations of teacher activities were pointed out by the interviewer. This aspect of the interview was one of the bases for discussion about change in the student teacher's thinking about science teaching. This collaborative interview allowed the novice teacher an opportunity to describe changes in thinking about teaching and learning in the novice's own words. Other information from survey results and informal conversations allowed deeper understanding of individual novice teachers' views about what teachers think about and do.

Personal constructs demonstrated for the student teachers themselves how they had changed in ways of categorizing the activities and events of teaching. In the fall, for example, one novice teacher distinguished activities that she stated were *activities which involve real teaching* from those that were *just administrative details*. In September, she placed the item "marking papers" in the category *just administrative details*. In February she placed the same item in the opposite pole of this category, *activities which involve real teaching*.

When this change was noted in the interview, the student teacher expressed surprise that there had been such a complete reversal in her thinking. She commented on how she believed her thinking had been transformed:

> I now see it [marking] as an important way of interacting with students on a daily basis. When I mark a paper, I'm making a comment to students, sometimes, I write one down. It's actually a way of communicating with them. I don't just correct it and give it back to them, I see what they did (laughs). I never thought of it that way when I got papers back myself, though. It's a way of talking to the students.

Personal Constructs as Collaborative Tools for Reflection

Personal constructs were used in the interviews as "tools for sensitive listening" (Mischel, 1964). The terms making up the constructs, contributed by the students themselves, helped to clarify original views about teaching and showed how these developed and changed from September to February. With the assistance of the interviewers, novice teachers reflected on growth and change in their views about teacher thought and action. Changes in the use of personal constructs served to direct attention to changes that occurred for student teachers during the methods and practicum experiences.

The great value of the theory in the collection of information is that it allows direction of attention to specific aspects of the experience. Student teachers fresh from their practical experience have often been asked the broad question, "What changes occurred for you in the student teaching experience?" This is a difficult question to answer. The novice has had so many experiences it may be difficult to decide which ones are most significant. In this study, reference to surveys and personal construct changes with the interviewer allowed students to discover how their thinking had changed in specific ways that were not immediately apparent to them.

Conceptualizing Changes in Students' Thinking About Science Teaching

The personal constructs proved to be a useful means of analyzing and reflecting on the new reality that was coming into existence for students. New images of teaching and learning emerged as students discussed the reasons for shifts in their constructs. New language structures emerged as students noted that occasionally their prior constructions no longer held, that new distinctions were now being made, and that new ways of construing experience were coming into existence.

The interviewer and the student teacher prepared statements to capture and convey the major areas of change noted in the students' thinking. The conversations were tape-recorded to preserve the students' own language. In addition to noting the changes, students were invited to identify the sources of change, the forces they considered at play, or to suggest possible reasons for the changes that had occurred.

The statements made in the collaborative interview were reviewed by the researcher for each individual case. Personal construct charts and changes were also reviewed. What was of prime interest in this review were the changes in thinking and acting that were apparent *to the student teachers themselves*, the ways that student teachers described awareness of transformations and changes in their own development. Categories were developed

to organize and consider these changes. The change categories were revised several times. Two assistants tested the categories by sorting statements from the collaborative interviews into the designated categories. The categories (listed below) delineate several ways that students themselves appear to understand their own growth and change.

Category 1: Awareness of Ideas That Have Been Simply Corrected or Old Dilemmas Which Have Been Resolved

The student teacher is aware that factual information about the realities of teaching and learning have been simply corrected. The student teacher may also be aware that he or she has resolved an old dilemma or question about teaching. Example statements are the following:

> Teachers work much harder than I thought they did.
>
> I thought that I would have more intimate contact and interaction with students in the classroom than I did. There is a much wider gap between students and teachers than I thought.
>
> I thought that teachers planned together much more than they do. There doesn't seem to be a need to follow curriculum as closely as I thought.
>
> I thought I would work much more with individuals than I did.
>
> Preparing lesson plans is not as important as I thought. *Thinking* about what we do is important, but the detail in the lesson plans is not so important.
>
> I was surprised that the kids did not perform nearly as well as I thought they would.
>
> I thought that students would be far more interested in science than they were. I'd have to honestly say that of my three grade nine classes, maybe one person showed a genuine interest.
>
> I expected that the methods course would have us doing lots of unit plans. Maybe it was naive, but I thought that we would be able to plan everything ahead of time that we would need to teach ... so we could use it later on.
>
> I thought that labs would be a "regular thing" in the methods class in order to learn different ways of doing labs.
>
> My thinking has changed on the amount of time after class that I need to spend with students.

Subcategory: An Old Dilemma is Resolved

The student teacher is aware that he or she has entered the classroom with a question in mind that is now resolved. Example statements are the following:

> Now I know how much fits into an hour, what you can reasonably expect will fit into an hour.
>
> I always wondered what teachers do after school. Now I know. They work with kids, they talk with kids, they plan and have meetings.

Category 2: Awareness of the Acquisition of New Technical Know-How

The student teacher is aware that he or she has acquired new technical skill, ability, or capacity. Example statements are the following:

I know now how to set up a lab and organize the kids to help me.

The idea of constructing an exam was novel to me before. I can put one together now that's a fair test of knowledge, I think.

I've discovered how to pace a lesson and organize my questions logically, so the idea is *developed* not just presented.

I learned to be clearer about my expectations. I learned to be consistent and follow through with kids. I tried to be friendly at first. Big mistake!

It's important to know the students and the background they come from. I'm learning to relate what is taught to their world.

I found that having students take notes was very helpful for discipline. Also, the students had trouble reading the text and taking notes, I wrote on the board and they would copy. (Notetaking moved from a teacher involvement activity to a student involvement activity on this student's construct list.)

I learned how to use questioning to involve students in lessons. I learned to ask open-ended questions and avoided closed-ended questions.

I learned to involve students in demonstrations.

I use overheads in class as a management technique. Every time I use the overhead the class is really quiet.

Category 3: Awareness of New Ways of Categorizing Experience or of Transformations in Categories Used to Describe Experiences

The student teacher is aware of transformations in his or her way of categorizing experiences. Shifts in thinking or the expansion of an idea are noted by the student teachers themselves. The student comments on the ways ideas have changed, been expanded, or been transformed. Example statements are the following.

In my mind, lesson plans are not just an outline for teaching. I see them now as a guide for involving students.

I thought of discipline as something separate from teaching before. It's a *part* of teaching, not something outside of it.

I now see all labs are creative experiences. Before I thought of them as conservative, routine learning events. They generate interest among the students.

I see marking as a daily interactive process. I thought it was just boring administratrivia before.

I've shifted from believing that activities should be "hands on" to believing in a more "traditional" approach. I'm sure my students would throw rocks if I gave them to them.

I've shifted in my view that one "works with teachers" a good amount of the time to believing that I'll work more with my own class, work with students! They are my responsibility.

I've discovered that notetaking by students is much less "student oriented" than I thought. It can be such a mindless activity for them.

I consider tests to be far more useful as a "planning tool" than as a "learning tool." I don't think kids are really interested in how they do on their tests. They are more for me.

I moved from considering "my involvement" in an activity to thinking more about students' involvement because the students like to take part in the demonstrations, for example.

I'm changing from "looking ahead" activities like planning to dealing with misunderstandings as they come up for students, to "doing things now."

I've shifted more towards activities for "students working together" from "me working with the students" because I find that they often ask one another about things first, before they ask me.

I see curriculum having a lot more flexibility. I thought it was very prescriptive before.

I discovered that I'm a science teacher but more than just that. I'm a lunch supervisor, home room teacher, holiday door decorator and dance chaperone.

Category 4: Changes in Self-Awareness. Knowledge of Self and of Self as a Teacher

The student teacher is aware of development in his or her own knowledge of self or self as teacher. Example statements are the following:

I saw myself as a student before—somewhat uninvolved. Now I see myself as a teacher.

I have a much clearer idea of what a teacher is and who students are. When I imagined myself in the classroom before, I pictured myself as silent. I feel more confident now. I have an idea about what being a teacher means.

I became more specific in my understanding of the meaning of flexibility in teaching and what it means to be motivated and active. I was not really thinking of myself as a teacher before.

I learn while I work with other teachers.

I've changed in my sense of confidence because I discovered that the kids like me. They related well to me. I didn't really learn so much about teaching strategies from my teacher. I also found out that I did a lot more content than I realized I did. This really helps me.

I've discovered that I'm really a content oriented person. Even more than I thought before.

I've found that my view of science teaching is different than that held by the profession. I have a more coherent comprehensive approach. Most teachers know very little about science or even how to teach it.

I have more confidence in my ability to deal with discipline.

I am very comfortable phoning up parents and speaking with them.

I'm more practical than idealistic now. Being in the classroom helped me mature, or gave me a slap in the face, I'm not sure which!

I found out that I have got to become more student-centered.

> I found that I cannot be as spontaneous as I want to be. In teaching you must stick to the curriculum, it's the burden on your shoulder.
>
> I've done more writing this year than I've ever done in my life. I've realized that it helps you to sort out things.
>
> I think that I'm really into getting the kids involved.
>
> I've become more aware of the differences between students and more aware of the need to be more tolerant and forgiving of students in general.

Category 5: New Dilemmas are Construed for Future Attention

The student teacher is aware that he or she is creating dilemmas to resolve in the future as he or she grows as a teacher. Example statements are the following:

> I have only recently begun to realize the importance of a philosophy of teaching and how it will help me to teach.
>
> I found that the curriculum can really tie you down. I'd like to focus on thinking, not just conveying content next round.
>
> I believe that if teachers' needs were catered to more, then the science program would improve. If administrators want change, then they will have to help teachers make that change.
>
> Language and communication skills should be emphasized more for them (students) to apply in everyday life and to enhance comprehension. We have to do more of this.
>
> I feel more comfortable about my knowledge of content but in the next practicum round I want to work on learning how to better work with individuals.
>
> I've come to the realization that education is not actually valued by government or even by some parents and that is something I would like to change.
>
> I've realized more how important it is to understand how students learn. There is a need for concrete activities.
>
> Most of what I'm going to do now will be oriented towards the curriculum. I imagine I'll be collecting lab materials which are curriculum based.
>
> I think that most teachers do not have a sufficient background themselves in Science Technology Societal issues. I'd like to see more professional development opportunities for teachers.

The Five Change Categories

The five change categories delineate not only the basic categories of change that occurred for students, but they are created to summarize the types of changes that students have seen in themselves as they develop a personal-social construction of the teaching role. Helping student teachers to acknowledge these changes using both systematic and informal discussion procedures enhanced the results both by providing focus for the discussions and contributing to the the credibility of the responses. Change category #3 is particularly rich in transformation statements as students demonstrate recognition of their prior views and how these have now changed.

Some Current Realities in the Task of Working Toward Collaborative Reflection in Teacher Education Programs

There are some important realities of the teaching methods courses on which we build our current programs. The first is that not only do teacher educators also attempt to prepare students for the task they will immediately undertake, the practicum experience, but teacher educators also attempt to prepare student teachers for their future role as professional educator. These demands are often seen to be competing for the student teacher's attention. Some teacher educators attempt to ease this dilemma by assisting student teachers to understand the value of possessing an educational philosophy. Teacher educators recognize that if one possesses a clearly articulated philosophy, the actions of teacher have purpose, depth, and meaning. For the most part, however, students experience great concern for survival and the competent completion of several weeks of practicum. Many students believe that we do not adequately address the latter concern in methods courses.

The change categories described hold potential for our helping to bridge the novice's need to survive the practicum while effectively reflecting on one's development as an educator. The change categories and student comments can be shared as case studies with student teachers to show the types of changes that typically occur as students develop an understanding of science teaching. Their use could allow students to understand what might be expected in the practicum setting and would encourage sharing of experiences in large and small group situations. New images and terminology help prepare students to articulate and clarify their own current viewpoints.

A second reality in teacher education programs is the fact that students are often members of a fairly large class. Experiences must be organized for students in large groups. Resources are not usually available to teacher educators for working with students individually or in small groups. The benefits of doing so have been clearly evident in our work in the study. Cooperating teachers and student teachers benefit from working together to clarify and articulate changes that occur during the teacher preparation experience. The research approach in the present study allowed student teachers *themselves* to articulate and affirm changes and transformations in their growth that they might not have been able to otherwise construct. The collaborators have been able to help students create their own personal meanings for experiences as they reflected on them together. In particular the personal constructs generated by the students served as powerful categorizations of experience that were useful in noting shifting and changing in their thinking.

Toward New Approaches and Possibilities

The research study is based on an assumption about personal-social change, that is, that our present interpretations of experience are subject to revision and replacement. Following this assumption, preservice teachers are more likely to accept new viewpoints if they can become aware of the personally meaningful implications of a new way of thinking. To do this, they must be aware of the personal-social construction of the teaching role and that it is a process of growth and change. Their own change can be seen in the context

of the culture of teaching. The effects of this reflection on the initiation into this new reality often requires revision and change in a student's viewpoint. A new view of the learner or approach to a laboratory experience is far more likely to be entertained when student teachers see the benefit of making the change in viewpoint for themselves. Greene (1973) has written: "No matter what degree of objectivity is attained, no matter how universally accepted are the rules, the final decision is up to the individual" (p. 253).

The Social Construction of the Science Teaching Role

The work gives important insight into our understanding of the variety ways that student teacher views and beliefs concerning student teaching change and are transformed in methods courses and practicum experiences. In the study, for example, the category *self-awareness* found a number of students discussing changes in their sense of self-confidence and in their sense of preparedness. The reasons given by student teachers for changes here were often the result of enhanced views of students' sense of ability and preparedness. In discussion, the students noted that this occurred simply from exposure to the student teaching experience. In several cases, students' stated levels of confidence and preparedness changed, in the mind of the student teacher, because of their realization of their own natural abilities. These students recognized in themselves a greater capacity to undertake the tasks of teaching than they realized prior to the experience. There are implications for sharing this information in the initiation of student teachers into the profession. The work suggests that the collaborative role of a caring guide, in the case of the research the research interviewer, is significant in assisting the student to articulate these changes. In the case of actual practice, the cooperating teacher or faculty representative could undertake a similar role. Several suggestions are made for building a collaborative, reflective interview into the teacher education program.

1. We can help student teachers establish some clear benchmarks at the beginning of their experience as a basis for reviewing and comparing their own development and growth. We can encourage them to consider their own development as teachers in the context of others' experience by sharing research findings such as those reported in this study as case studies.
2. We can find new ways to build on those aspects of the student teaching experience that student teachers have found most beneficial in reflection on their growth and change. In the course of this study, novice teachers identified several experiences as contributing positively to their development. Some of these were: viewing videotapes showing various science teaching approaches and techniques; and taking a course on content reading across the curriculum.
3. Teacher educators can effectively develop new approaches for working with collaborating professionals such as cooperating teachers and faculty consultants. Involvement of teachers in this special collaborating role is a valuable professional development opportunity. We can build guidelines for reflective interviews that suggest how collaborators in the field might

share in the uncovering of changes in student teacher thinking to understand growth and change in the student teacher's development together.

An approach that strives to make student teachers aware of changes in their growth also asks that we consider our own growth as educators. Westerhoff (1987) writes:

> We are at our best when we make our lives and our search for meaning available as a resource for another's learning. To be a teacher means more than to be a professional who possesses knowledge and skills. It is to have the courage to enter into a common search with others. (p. 193)

Bolin (1987) comments on the renewing power of the teacher's work with students. He states that it is our role to be creators as well as bearers of meaning. But we must also develop ways of renewing our practice.

> Our work is to empower students to find their own personal meaning. But we cannot do that year after year without attending to our own meaning making and empowerment. (p. 219)

We must go on to sustain renewal. Such renewal is possible with new views of teacher education programs. John Goodlad, whose quote serves to introduce this article, comments that new reform in teacher education requires a "simultaneous renewal" of both teacher preparation programs and the public schools. Our attempts to help preservice teachers grow and change must mean our own renewed vision of what it means to be a teacher. This vision will guide us in the ways that we come to know student teachers and how we help them to know themselves as teachers.

References

Bolin, F.S. (1987). Teaching as a self-renewing vocation. In F.S. Bolin & J. McConnel Falk (Eds.), *Teacher renewal: Professional issues, personal choices* (pp. 217-230). New York: Teachers College Press.

Calderhead, J. (1987). *Exploring teachers' thinking*. London: Cassell.

Calderhead, J. (1989). Reflective teaching and teacher education. *Teaching & Teacher Education, 5*(1), 43-51.

Erickson, G.L. (1987). *Constructivist epistemology and the professional development of teachers*. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Greene, M. (1973). *Teacher as stranger: Educational philosophy for the modern age*. Belmont, CA: Wadsworth.

Grimmett, P.P. (1988). The nature of reflection and Schön's conception in perspective. In P.P. Grimmett & G. Erickson (Eds.), *Reflection in teacher education* (pp. 5-15). Vancouver, BC: Pacific Education Press, University of British Columbia; New York: Teachers College Press.

Grimmett, P.P., & Erickson, G. (Eds.). (1988). *Reflection in teacher education*. Vancouver, BC: Pacific Education Press, University of British Columbia; New York: Teachers College Press.

Hewson, P.W., & Hewson, M.G.A.B. (1988). *Analysis and use of a task for identifying conceptions of teaching science*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Kelly, G. (1969). The strategy of psychological research. In B. Maher (Ed.), *Clinical psychology and personality: The collected papers of George Kelly*. New York: Wiley.

Kelly, G. (1970). Behavior as an experiment. In D. Bannister (Ed.), *Perspectives in personal construct theory*. New York: Academic Press.

Mischel, T. (1964). Personal constructs, rules, and the logic of clinical activity. *Psychological Review, 71*(3), 180-192.

Osborne, R., & Gilbert, J. (1985). Some issues of theory in science education. In R. Osborne & J. Gilbert (Eds.), *Some issues of theory in science education*. Hamilton, NZ: Science Education Research Unit, University of Waikato.

Schön, D.A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Schön, D.A. (1988). *Coaching reflective thinking.* In P.P. Grimmett & G. Erickson (Eds.), *Reflection in teacher education* (pp. 19-29). Vancouver, BC: Pacific Education Press, University of British Columbia; New York: Teachers College Press.

Shapiro, B. (1988, April). *Learning about thinking, thinking about learning.* Paper presented to the National Association of Research in Science Teaching, San Diego.

Watkins, B.T. (1990, February 28). Reform of primary and secondary education linked to renewal of teacher training, education-school deans are told. *Chronicle of Higher Education*, pp. A17-18.

Westerhoff, J.H. (1987). The teacher as pilgrim. In F.S. Bolin & J. McConnel Falk (Eds.), *Teacher renewal: Professional issues, personal choices* (pp. 190-201). New York: Teachers College Press.

Appendix A
Item List A
What Teachers Think and Do
Student Teachers' Ideas

1. Marking student papers.
2. Using materials other than the textbook to plan lessons.
3. Doing experiments with the class.
4. Maintaining discipline.
5. Finding out what the kids already know before I teach.
6. Getting students involved with and committed to preserving the environment.
7. Encouraging students to do well.
8. Using analogies to explain ideas.
9. Doing demonstrations.
10. Working together with teachers.
11. Locating the laboratory equipment I need.
12. Having students take notes.
13. Planning lessons.
14. Going on field trips.
15. Lecturing to students.
16. Being flexible. Deviating from the curriculum. Having students pursue topics of their own interest.
17. Dissecting frogs.
18. Learning about the curriculum and curriculum changes.
19. Working with students individually.
20. Helping a student to overcome a difficult experience in his or her personal life.

Appendix B
Item List B
What Teachers Think About and Do
Practicing Teachers' Statements

1. I look for materials and lab equipment. Consider any supply problems that I might have.
2. We are often doing readings. Doing these together as a group or individually, then answering questions from the readings.
3. Having students do library research or smaller reports such as newspaper, magazine research or perhaps a visit to a doctor.

4. As I teach, I move around the class, moving from one group to another or from person to person answering questions and locating equipment for them, or getting them text resources.
5. In the high school I find that I 'wander around in the curriculum a fair amount'—diverge for several periods to work on items of interest, like cold fusion. When I can't continue in class I might meet with them or a small group individually at a prescribed time.
6. I deal with current events and the technological applications of content, like cold fusion when the kids are interested.
7. I try to understand why the kids are there. Like, are they going on in science or are they going into careers right out of high school? Then I know what strikes them/what hits them. That affects their understanding.
8. I consider the wide intellectual range of kids in my class and spend a lot of time trying to encourage everyone to do the best they can.
9. There's a constant need to maintain discipline. I'm constantly moving around the classroom making sure people are on track.
10. I try to give the students a fair amount of feedback on tests and reports.
11. The first thing I do is review the curriculum materials. I check for the organization of things, the content.
12. I look around to see what kinds of activities are available to teach a concept.
13. I review the text. I think, for example, is it too young for the kids? One course I taught was demeaning for the kids, it was too young for them. I look around for alternative materials and change and adapt when necessary.
14. I start where the students are. I try to find out where they are by getting in touch with their current level of understanding.
15. I consider the developmental level of the class as a whole, and what that means, for example, to the non-academic student.
16. In the classroom I think about whether I really need to duplicate material or whether or not I can show an activity on the board instead, or do a demonstration to reduce costs. Too much paper is given out in classrooms.
17. I spend time getting down the rules of discipline in the classroom.
18. I involve the kids in pencil and paper activities. In these, they answer questions about the material we have covered.
19. I have the kids go back and think about what they have done in the previous lessons. Relate the lesson to, for example, the one from a week ago.
20. When giving an introductory lesson, I tell the kids what they are going to be working on rather than keeping it a secret.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 2, June 1991, 133-140*

DAHLIA BECK
University of Calgary

and

KERRY BLACK
University Elementary School, Calgary

Redefining Research Relationships: Two Heads Are Better Than One

This article describes a process of collaborative research between a primary grade school teacher and a researcher in the context of a case study of teacher planning. The paper explores the dual and interrelated courses of mutual meaning construction in both planning and collaboration during the coparticipants' joint work. The study demonstrates how educational research definitions and hence research options can be broadened and diversified, affording new possibilities to generate knowledge and enrich our understanding of the field.

Teacher planning has traditionally been described as a one-dimensional linear process, whereby the practitioner inserts the "appropriate" teaching skills (prescribed by the curriculum) into objective terms (rationale, behavioral objectives, method, and so on), systematically implementing the skills in the classroom. Within the framework of this model, the teacher is regarded as a "conduit" (Clandinin & Connelly, in press) between the prescribed curriculum and the actual realization of lessons, and not as a professional who is capable of constructing knowledge. However, recent research on the teacher thought domain (Clark & Peterson 1986; Clark & Yinger, 1980) has shifted the focus from the product of planning to the planning process, investigating and uncovering its complexities. The teacher thought domain has accordingly been described as encompassing the constantly interacting teacher's theories and beliefs, teacher planning—the preactive and postactive stages, and teacher's interactive thoughts and decisions (Clark & Peterson, 1986). The practitioner, according to this model, is seen as a central and integral part of the planning process and is no longer viewed solely as a transmitter of knowledge, but as embodying and generating knowledge. Thus the knower and the knowing are perceived to be inherently connected.

Our joint project began as an investigation of a portion of teaching and the teacher thought domain that forms part of the foundation of the classroom context, namely, planning for teaching—the preactive and postactive stages.

Dahlia Beck is a doctoral candidate in the Department of Curriculum and Instruction. Her research interests focus on teacher education and second language instruction.

Kerry Black is a primary school teacher with research interests in school-university collaboration and teacher knowledge.

Each of us was interested in planning from a different perspective. Kerry, the teacher, wanted to find out how she knows what she knows about teaching and what makes her plan to teach the ways she does. She felt that teacher knowledge could be tapped by observing a teacher plan and discussing the process, and product, with another. The researcher, Dahlia, wanted to focus on the theoretical underpinnings of teacher knowledge, on how teachers generate and use their knowledge, as manifested in their planning to teach. As collaborating partners we were hoping to gain insight into planning and practice together. Because each of us was approaching the study with a different personal and experiential narrative, we felt that a mutually constructed vision afforded new possibilities that, though we could not anticipate it, we were anxious to explore.

Kerry has taught in the public school system, at the preschool and elementary level, for 12 years and brought her ongoing classroom experience, and queries thereon, to the investigation. She was about to try new math and reading programs with her students and felt the need for feedback, for exchanging ideas with another practitioner, with someone who would have a different, preferably critical, perspective on her ideas and practice. Kerry was also doing graduate work in teacher education. Dahlia, who taught English as a foreign language for 15 years at the junior- and senior-high school levels, as well as the tertiary level, is a full-time graduate student majoring in second-language teacher education. She focuses her scholastic energy on researching the theoretical and practical aspects of the concept of teaching and on the relationship between teacher practice and teacher knowledge.

We met in a graduate course that centered on teacher practice as explained and made sense of by teachers through their stories and metaphors, a framework developed by Clandinin (1986, 1990), Clandinin and Connelly (1986, 1988), and Connelly and Clandinin (1988, 1989) to explore teacher knowledge. Through class discussions we discovered our mutual interest in the planning process, which led us to consider working on a joint project, thinking that collaboration afforded a possibility to study planning from multiple perspectives simultaneously, something neither of us had done before.

Following a mutual declaration of intent, we were faced with an ethical issue: Because Kerry is a practicing teacher with the Calgary Board of Education, and as we intended to voice our findings publicly, Kerry's anonymity had to be preserved to fit within the ethical guidelines for university research. Despite our understanding that the anonymity was a protection measure, we were dissatisfied with the fact that should a published paper result from our joint work, only Dahlia's real name could appear on it while Kerry's would be concealed behind a pseudonym. That necessary silencing, we felt, violated the true collaborative nature of our work.

Although inexperienced with collaborating on a research project, we nevertheless knew that we wanted the study itself to be a form of dialogue between us, where each would have as much a say as the other on each one's work and on each one's interpretations of the joint endeavor. The conversation approach we decided to adopt is consonant with Yinger's (1987), who elaborately uses conversation as a metaphor for practice. "First," he main-

tains, "conversation refers to the means by which social practices are conducted ... Second, the conversational metaphor expresses the multifaceted, give-and-take nature of human thought." And third, as he examines the etymology of conversation, Yinger argues that it "involves an entering into and living with a context and its participants" (p. 11). As such, he regards conversation as a means of interaction and a way of thinking, as a type of relationship with one's surrounding. The conversation metaphor, we felt, would help ensure fairness and authenticity in the researcher's representation of the teacher's work, as well as aid in telling the story of the teacher and her work in a light that was compatible with her own image of herself. The phenomenology of the teacher, that is, the meanings of planning for her, from her point of view and frame of reference, we thought, could be best captured via conversation. For us that was the essence of the *co-labor* notion.

To enhance the oral dialogue we had agreed to keep journals where we would relate to our joint project and read them during the period of our work together. Kerry had had experience with journal keeping, a process she found to be a powerful reflective tool. Dahlia had none. We fell back on Kerry's experience, as well as on Connelly and Clandinin (1988) for whom journals are "ongoing records of practices and reflections on those practices" (pp. 34-35). Indeed, the journal served as a useful refresher of interpretations and deliberations that captured, at the time of writing, the truths of the moment concerning the day-to-dayness of classroom events and our joint work, assisting us, when studied together with the grounded data later on, in further understanding the process under investigation. The journal proved to be a powerful inquiry device, providing us an additional lens through which the data could be viewed (triangulation), its forte being its immediacy and authenticity. As there were intervals between our sessions, each of us felt it necessary to record those things that came to mind and happened between meetings as a means of memory. This became a link, a form of reflection with the self, and each other. The journal entries recorded the perpetual dilemmas of daily practice that are part of the detail of lived experiences. These day-to-day nuances formed a meaningful way to examine practice, planning, our relationship, and roles in the research.

How many times should we meet? we wondered, and decided tentatively to meet regularly for a period of at least two weeks, or alternatively until we felt we had covered enough ground to warrant a halt in data collection. The data, then, were going to lead us rather than vice versa. Following the latter route, we met seven times during a three-week period for working sessions that lasted one to three hours each, all of which were recorded and later transcribed. We had agreed that Dahlia would analyze and interpret the transcribed data and write a narrative account of them, by which was meant, following Connelly and Clandinin (1988, 1990), that she would relate how the teacher and the researcher, separately and together, made meaning of their experience. She would be the person who, according to Connelly and Clandinin (1990), "is at once ... engaged in living, telling, retelling and reliving stories," thus "setting and resetting temporal and social/cultural horizons" (p. 8). Dahlia would then hand the account over to Kerry who would respond

to it in writing. This process was to continue until both of us were satisfied with the representations in the mutual account, in which the teacher's and the researcher's narratives would become, in part, "a shared narrative construction and reconstruction through the inquiry" (p. 8).

We adhered to the above plan. However, following the spirit of conversation, the narrative accounts that we exchanged were in letter form, which was found to be, as Connelly and Clandinin (1988) point out, "valuable in aiding reflection [in the] ongoing dialogue between two teachers" (p. 48). A letter entails a mutual obligation and encapsulates the collaborative stance we chose to adopt. A letter is not descriptive and public oriented. It is directed to a particular person whose genuine reaction to the content the writer seeks, for it concerns the writer and addressee intimately, and the warrants and truths established in it have to be confirmed, validated, or at least addressed. A letter's asset, then, lies also in the response it triggers.

The research project began when Kerry was about to start planning and teaching new units; we thought that she would do the planning and Dahlia would observe her doing it. Kerry was intrigued by the notion of being observed while planning. "What would there be to see? My planning is largely mental, and the written plans would have already been processed," she contemplated aloud. Kerry pointed out at our very first meeting, before the collection of data began, that she had a vision of where she was heading with her students before she actually wrote anything down on paper. Dahlia, on the other hand, expected Kerry to progress in her planning in a traditional, linear fashion, assuming she would do it all, that is, have a finished product, prior to implementation. But Kerry did not plan like that. In fact, Dahlia's vision ran contrary to Kerry's understanding and conceptualization of planning. The relationship between planning and teaching for Kerry is dialectic, whereby each part instructs and builds on the other. Kerry's pre- and postactive plans depend largely on the interactive stage. For her, the preactive stage is a vision of a plan. However, it does not exist without the students who give the plan life, meaning, and direction at the interactive phase. The students' responses to the lesson determine the next day's lesson, and so on, until the goal set by the original vision is realized. Thus the students connect vision and reality into an experiential whole.

Thus Kerry's conceptualization and execution of planning was a major force in shaping the texture of our conversation. Instead of the anticipated long (teacher) planning silences, interrupted occasionally by (researcher) questions, our joint sessions became intense and animated by Kerry's stories of, and Dahlia's queries about, that day's interactive phase that illuminated and guided the next step of Kerry's planning. In other words, the concrete teaching plans, as opposed to the vision and the mental plans, were generated parallel to or after the interactive phase, bearing the latter in mind and responding to it, as well as to the reasoning and inquiry processes that permeated our joint sessions. To us, this was an instance of Schön's (1983) knowledge-in-action that we came to view as a distinct kind of knowledge, different from theoretical knowledge, yet in constant interaction with it.

Our respective roles in the joint process, apart from the labels "teacher," "researcher," and "collaborators," was a terrain to be broken through and

carved as we proceeded, together and individually. For Kerry, who had co- and team-taught, Dahlia's presence was a treat, for she had a person with whom she could play the math games she hoped to use with her class. She was looking forward to Dahlia's asking questions, which, Kerry felt, might help her clarify her thoughts and her rationale as pertaining to planning. Dahlia, who had storied her role as the researcher along traditional lines, perhaps because she was in the grip of a myth regarding researcher/informant roles, thought she was going to be an observer-interviewer, that is, Kerry, the teacher, would be working on her planning and thinking aloud (for the tape-recorder), and Dahlia, in order to minimize interference and background noise, would keep silent and jot down descriptions of the circumstances. Later she would interview Kerry about her planning. The clarity of this vision became blurred as soon as the first working session began. Rather than one interviewing another in a neat, orderly fashion, we conversed about our practices, conducting animated discussions about issues related to Kerry's planning that arose spontaneously during the planning sessions. It was messy. Thus Dahlia's brief role as an observer-interviewer ended and she became a participant-observer, a role that was also short-lived. As our relationship developed through our intense encounters, writing journals, and exchanging letters, the label participant-observer seemed inadequate, for it did not capture the dynamics of our alliance. The sociologist Reinharz and her colleagues (Belenky, Clinchy, Goldberger, & Tarule, 1986) came to our aid with a phrase they coined—temporary affiliation—which they feel captures the human mutuality that characterizes the relationship between researchers and informants. Hence we became temporary affiliates.

In retrospect we have come to regard our conversation/research as a spiral web comprising experiences and discourses that together, as they encounter and interact, constitute a meaningful whole. By experiences and discourses we refer, for example, to Kerry's teaching, the stories about it, and about the students that Dahlia experienced vicariously; the journals that we composed, exchanged, and responded to; our recorded conversations and the following narrative accounts; related readings and impressions thereof that we exchanged; and our university course work. We shared our individual discourses, which were simultaneously modified and recast by that sharing, thus expanding the web constantly. The discourses and lived experiences formed and shaped our work and relationship; they in turn fashioned and framed our individual and joint experiences and discourses.

With hindsight we now realize that both of us were wary at the beginning of the joint research, not quite knowing if and how the project would work out, especially as we had no firm guidelines to follow. We carved and forged new territory and language. Yet we felt that a productive research study would depend greatly on the mindful relationship we formed. Our common goal and genuine mutual interests bonded us and shaped the nonthreatening circumstances that evolved, enabling us to tackle our project. Once this process took off, it flowed vigorously. We should emphasize that this process very much resembles that of Kerry's planning. That is to say, we entered the partnership with visions and mental plans that, along the way, encountered other realities, other experiences and discourses, that had to be taken into

account. These multiple encounters affected our work and knowledge, specifically the personal practical knowledge (Clandinin, 1986) that we, the practitioner and the researcher, embody and employ as we practice our respective crafts and as we collaborate. Similarly, Kerry's planning is the outcome of multiple encounters of a variety of elements that carve and fashion it. During this dynamic process, as theory and practice come in contact, the original visions and mental plans are transformed and new ones, novel relationships and stories, and fresh narratives are constructed.

Collaboration, we discovered, is a powerful and instrumental device in uncovering and tackling individual values and beliefs, biases, and preconceived notions. As Peshkin (1988) observes, "Knowingly or not, I think we all are ... in the subjective underbrush of our own research experience" (p. 20). In collaborative circumstances, where each party can act as a mirror and a challenger, the subjective underbrush can be carefully examined through inquiry, reflection, and critical thinking, thereby generating knowledge.

The collaboration experience and its investigation, as we understand it, can bring about the expansion of method and forms of representation that, according to Eisner (1988), are all partial, and "because they are partial, they limit, as well as illuminate what through them we are able to experience." Through the expansion of method, in our case—collaboration in educational research—"our politics," Eisner maintains, "will become a liberating force for both understanding and enhancing the educational process" (pp. 19-20).

The joint endeavor is our statement about the optative relationship between educational practice and research. Traditionally, the relationship is top down; that is, the teacher and researcher are distant and separated hierarchically and view each other, respectively, as a receptacle to be filled (who will, in turn, transfer the "merchandise" to the students) and an expert, who embodies and delivers the necessary "goods." Accordingly, the researcher has the knowledge, and hence the power, while the teacher, who is a mere technical object in the chain of knowledge transfer (education), is not endowed with reflecting and reasoning capacities, nor is he or she viewed as an autonomous professional who embodies and generates knowledge. Such a stance invalidates and silences the teacher, disfranchising his or her knowledge. It also represents knowledge as an objective, complete, and remote entity that is intrinsically impenetrable, hence static and conservative. Our conceptualization of the teacher-researcher liaison is similar to Polanyi's (1964) concept of indwelling that refers to a sharing and satisfying of intellectual and convivial passion by two equals who "must belong to each other by participating in a reverence for a common superior knowledge" (p. 378). Indwelling characterizes the interrelationship between several aspects on which, Polanyi argues, cultural and ritual fellowship is based. They include the sharing of convictions and of a fellowship and cooperation (p. 212). Indwelling recapitulates our joint work in which the teacher and researcher were equal partners trying to make sense of, and converge on, the area of inquiry together from multiple perspectives.

The mutual quality of our associated research stems from, and is indicative of, a moral stand to which we are committed, of which a central element is communicative competence. It is an instance of Connelly and Clandinin's

(1990) narrative as a research method that involves sharing, equality, caring, and mutual purpose and intention. Collaboration, as captured by the narrative stance toward educational research, empowers the partners and helps them gain a voice that Britzman (1989), for example, regards as a central construct in curriculum, teaching, and educational research. "Restoring voice requires acknowledging relationships: the individual's relationship to the meaning of her/his experience ... and the individual's relationship to others," she argues (p. 146).

We understand the notions of collaboration, voice, and communicative competence as interpreted and realized in our research to be compatible with the principles underlying democracy. Democracy, in its ideal conceptualization, is a cooperative achievement of participation, dialogue, and creation, oriented to reaching shared understandings. Yet, as Habermas (1984) points out, coming to an understanding is not "an empirical event that causes de facto agreement; it is a process of *mutually convincing one another* in which the actions of participants are coordinated on the basis of motivation by reasons" (our emphasis, p. 392). A communicatively achieved agreement rests on common convictions and has a rational basis. It is a cooperative process of interpretation in which, Habermas contends, "no participant has a monopoly on correct interpretation" (p. 392). Moreover, resonating with Polanyi's (1964) notion of indwelling and Clandinin's (1989) narrative, Habermas (1984) argues that the communicative and interpretive task for all parties "consists in incorporating the other's interpretation of the situation into one's own in such a way that in the revised version 'his' external world and 'my' world can—against the background of 'our' lifeworld—be relativized in relation to 'the' world, and the divergent situation definitions can be brought to coincide sufficiently" (p. 100). Similarly, in the interpretive process of the narrative, "meanings are not fixed or given," Belsey (1980) claims, "but are released" (p. 20) negotiated and constructed.

A variety of discourses were brought together in our research, by which new and more coherent discourses were produced. Collaboration via the conversation metaphor as we experienced it enhances variety. Hence the resulting narrative, following Belsey (1980), "is not restricted to a single, harmonious and authoritative interpretation. Instead it becomes *plural* ... an object of work to produce meaning" (p. 104). Furthermore, the new discourse that emerges constitutes new knowledge, thus acting as a generative theory (Tom, 1984), adding to our individual and collective generative metaphors and repertoire (Schön 1983).

Collaboration between teacher and researcher as we perceive it promotes engagement in critical thinking and practice, thus constructing knowledge and cultivating the participants' capacity to structure and appropriate their knowledge and understandings, as well as to develop a language with which to do so. We regard these processes as conducive to change, allowing for pedagogic possibilities.

References

Belenky, M.F., Clinchy, B.M., Goldberger, N.R., & Tarule, J.M. (1986). *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.

Belsey, C. (1980). *Critical practice*. London: Methuen.

Britzman, D.P. (1989). Who has the floor? Curriculum, teaching, and the English student teacher's struggle for voice. *Curriculum Inquiry, 19*(2), 143-162.

Clandinin, D.J. (1986). *Classroom practice: Teacher images in action*. London: Falmer.

Clandinin, D.J. (1989). Developing rhythm in teaching: The narrative study of a beginning teacher's personal practical knowledge of classrooms. *Curriculum Inquiry, 19*(2), 121-141.

Clandinin, D.J., & Connelly, F.M. (1986). Rhythms in teaching: The narrative study of teachers' personal practical knowledge of classrooms. *Teaching and Teacher Education, 2*(4), 377-387.

Clandinin, D.J., & Connelly, F.M. (1988, April). *Narrative, experience and the study of curriculum*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Clandinin, D.J., & Connelly, F.M. (in press). Teacher as curriculum maker. In P.W. Jackson (Ed.), *Handbook of research on curriculum*. A project of the American Educational Research Association.

Clark, C.M., & Peterson, P.L. (1986). Teachers' thought processes. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255-296). New York: Macmillan.

Clark, C.M., & Yinger, R.J. (1980). *The hidden world of teaching*. East Lansing, MI: Institute for Research on Teaching, Michigan State University.

Connelly, F.M., & Clandinin, D.J. (1988). *Teachers as curriculum planners: Narratives of experience*. New York: Teachers College Press.

Connelly, F.M., & Clandinin, D.J. (1990). Stories of experience and narrative inquiry. *Educational Researcher, 19*(5), 2-14.

Eisner, E.W. (1988, June/July). The primacy of experience and the politics of method. *Educational Researcher*, 15-20.

Habermas, J. (1984). *The theory of communicative action, vol. 1, Reason and the rationalization of society* (T. McCarthy, Trans.). Boston, MA: Beacon Press.

Peshkin, A. (1988, October). In search of subjectivity—One's own. *Educational Researcher*, 17-21.

Polanyi, M. (1964). *Personal knowledge: Towards a post-critical philosophy*. New York: Harper & Row.

Schön, D.A. (1983). *The reflective practitioner*. New York: Basic Books.

Tom, A.R. (1984). *Teaching as a moral craft*. New York: Longman.

Yinger, R.J. (1987, October). *The conversation of practice*. Paper prepared for the Reflective Inquiry Conference, University of Houston, October.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 2, June 1991, 141-156*

CHUCK CHAMBERLIN

and

JAN VALLANCE
University of Alberta

Reflections on a Collaborative School-Based Teacher Education Project

A shared concern for teacher education brought together staff from the University of Alberta, Faculty of Education and Edmonton school principals and teachers who designed and implemented an in-school preparation course for university students. Preplanning between staff from three schools and a group of university professors resulted in the placement of 68 undergraduate third-year education students into classroom experiences. The students were engaged in two university curriculum and instruction courses of nine hours per week. One course combined social studies, music, and language arts and the other, movement, art, math, and science. This article interprets student journals and interviews with students, school staff, and professors in order to illuminate the experience of the participants. The collaborative nature of the experience, the interactions, and the reflection resulting from this university-school program are discussed. Further examination of the commitment to such an undertaking, the time required for collaborative staff procedures, and reflection time between students and professionals are considered.

Introduction

During September through December 1987 faculty members from the Department of Elementary Education at the University of Alberta and principals and teachers from three elementary schools in the Edmonton Public School system met to plan a collaborative teacher education program. They held a mutual interest in exploring a different vision for university teacher education courses. Through sharing experience and perspectives, they searched for the possibilities to be found in centering the teacher education courses more within elementary school sites. The purpose was to design jointly a program to provide more opportunity for students to learn teaching by teaching and reflection. Teachers and professors collaborated in designing and carrying out such a reconceived approach to teacher education.

This project was to involve 68 undergraduate university Bachelor of Education students in school classroom experiences for a half-day each week, which would coincide with their module courses in music, language arts, and

Chuck Chamberlin is a professor in the Department of Elementary Education. His recent research has examined the effects of teacher education on teachers' beliefs and ideologies.

Jan Vallance is an associate professor in the Department of Elementary Education. Her collaborative research into the meaning of children's creative dance examined teachers' beliefs and personal credo.

social studies methods or math, science, art, and movement education methods. Assignments in these courses would include student planning, teaching, and evaluating activities undertaken with teachers and children in a school setting. Opportunities for observation, assisting in the teacher's lessons, and teaching their own lessons would be provided. Students' journals would be used to record their observations and reflections on the work with children in all of these situations. To establish a nonintimidating exposure to teaching and a free interaction between participants, student journals would not be graded. To promote discussion, reflection, and to provide emotional support, pairs or small groups of students would work together on assignments. The plans thus developed included on-campus and in-school experiences as a nine-credit course (normally nine hours per week).

The Roots of the Problem

Teacher educators are challenged by studies such as Fuller and Bown (1975), Feiman-Nemser and Floden (1986) to understand how beginning teachers think and act. These studies indicate that teachers' early experience as pupils influence their expectations about teaching and also their teaching behavior. Those entering teacher preparation have already had 12 or so years of interactions with experienced teachers, and this has given them strong messages about what teachers do. Feiman-Nemser and Floden (1986) describe the socializing power of the university as weak compared with the competing norms of schools and refer to Zeichner and Tabachnick (1981) who state that the effects of university socialization are often "washed out" by the strong influence of the beginning teachers' own schooling experience (p. 7).

Further research has pointed to the continuing nature of socialization into teaching perspectives. Like Fuller and Bown, Craig (1984) came to the realization that teacher development is a continual process of preservice, beginning teaching, and inservice, following steps or stages in one inclusive continuum. "Currently they are recognized more and more as developmental, integral, and complementary parts of a rational whole" (Greenberg in Craig, 1984, p. 259).

Other studies suggest that teachers' perspectives are influenced by preservice education, inservice education, and informal peer socializing. Lacey (1977) found that some teachers conformed to the new expectations and practices; others, not convinced that the practices were for the best, still complied, while some beginning teachers consciously tried to effect change or reform. Fuller and Bown (1975) suggest that as education students move from methods courses to student teaching they "change from a humanistic to a custodial approach, stressing bureaucratic order and control.... These changes may reflect a shift during student teaching toward the prevailing ethos of the public school" (pp. 41-42). Tabachnick (1980) reports a study of interns who changed from wanting warm relationships with students to desiring cooler, more distant relationships. He concludes that "public school staffs have an interest in legitimating their actions and keeping things the way they are" (p. 133). Iannacconne's earlier research (1963) using interns' log books led to much the same conclusion. These studies suggest the need to

examine the ways teachers' and students' perspectives change as they interact with university and school norms and expectations.

Newberry (1979) found that beginning teachers were usually influenced by teachers who taught the same grade level and had similar teaching styles. Further, Tabachnick and Zeichner (1985) suggest that perspectives developed during preservice teacher education will be influenced by peer socializing efforts during the first year of teaching and will only persist in individuals who have a mixture of skill in building support from parents and students, and their own persistence. These researchers "support an interactive view of teacher socialization in which individual intent and institutional constraint both play a role in affecting a beginning teacher's entry into teaching role" (p. 10).

The need for collaboration of teachers and professors in planning teacher education programs emerged from two studies during 1986-1987 of recent graduates of the BEd program in elementary education at the University of Alberta. Miklos, Green, and Conklin (1987) found that most graduates from the teacher education programs at universities in Edmonton, Calgary, and Lethbridge "were at least mildly dissatisfied with the extent to which the program had met their expectations" (p. 136). Concerns focused on lack of emphasis on teaching skills (p. 147), too much emphasis on theory (p. 149), redundancy (p. 151), inadequate practicum (p. 151), and poor teaching models (p. 154). Teacher graduates interviewed suggested: "Whenever possible, I think that what is being taught should be pulled into real-life situations. As much as possible, get the students in the university in contact with the kids" (p. 160). The in-school experience of practice was seen as of most value, as "teachers rated this as the most important component of their preservice preparation program by a substantial margin" (p. 180), and asked that courses "be made more practical and relevant to the work of a teacher and less general and conceptual" (p. 161). This could be facilitated by "holding seminar-type classes with experienced teachers" (p. 162).

Blakey et al. (1989) used case studies of 14 recent University of Alberta graduates to examine the factors affecting teacher decision making and the perspective underlying those decisions. Again, teachers reported their BEd university based classes "ranked far behind the practicum (student teaching) component as the most valuable aspect of undergraduate study" (p. 21). One teacher stated, "I think I learned more in my practicum than I've learned in any of those topics or subjects or courses. You learn so much by just getting the practical experience" (p. 32). Recommendations suggested that the teacher education program include: "holding classes in public schools for several weeks to develop a feel for life in the school; providing in-class experiences with children so instructor and students experience working with children.... and teams of university professors and practicing teachers should be organized to offer courses in school settings" (p. 111). Such plans could help to meet the needs of teachers who stated, "I would have learned more by being out in different schools and kindergartens and ... just watching kids.... More university classes should be in the schools. Like, really find out what's happening. So I would say a lot more time [should be] spent in school" (p. 112). Blakey et al. felt that the connection between theory and practice

can only be made "when there is a close connection in time between the two" (p. 112). Courses that had some connected school experience and the opportunity to relate theory to specific children seemed to make the ideas more meaningful and increase the likelihood of later application in teaching.

The present study seeks to understand how preservice teacher educators may assist students in consciously evolving a teaching perspective through concurrent university courses and school experiences.

The Development of the Project

In September 1987, several University of Alberta professors began meeting with principals from three elementary schools in Edmonton to explore possible ways of building a team of teachers, principals, professors, and elementary education students who would develop a collaborative school-based teacher education program. The result of three months of planning was the placement of 68 education students in three schools for a half-day each week during their curriculum and instruction module course. Teachers and undergraduates would plan activities for children, try out methods course theories, and the students would reflect on those experiences in journal writing. For example, in their language arts course, students studied the use of drama, including an all-day workshop held in one of the schools. Several of the cooperating teachers took part in a Saturday workshop and later worked with students who tried out variations in the use of role playing, dramatizing, and tableaux in their classrooms. Students then used their journals to reflect on this classroom experience and the language theory they had learned.

In movement education, students worked in small groups assigned to one classroom. Together with their professor they planned and taught their class activities in gymnastics, games, and dance. Each student group followed their children from gymnasium to classroom where varied teaching responsibilities were provided to them by the class teacher. Over time, teaching staff and students made more specific requests and offers to undertake responsibilities such as planning a class orienteering experience, assisting in library research and activity centres. Throughout the in-school experience with the children, students reflected on their personal as well as group teaching and the interactions with the class. Journal entries provided the major data source for this study. Student voices were given the dominant place in the descriptions, analyses, and assessments that form this report.

Major Themes Emerging from the Study

The findings of the research team have been compiled from the journals and interviews undertaken by and with the undergraduate students and the interviews with the teachers and professors engaged in the collaborative school-based teacher education project underway during the January to April 1988 university term. The results are summarized under the three themes that emerged: (a) collaboration and commitment, (b) interaction, and (c) reflection. In each theme emphasis is placed on the voices of the students. The voices of the teachers and professors are also heard along with the interpretation of the researcher-participants.

Collaboration and Commitment

The "we" attitude. Students entered the school scene in groups of two, four, or seven depending on the number of staff members participating and the curriculum and instruction module involved. Recognition of coinvolvement with peers permeated much of students' writing. "Everyone in our group has a part ... This was the first time we had a chance to actually preplan together." "We built a trust relationship." "we joined in." "Our team worked well together." From these comments it appears that students realized a sense of security and support through peer involvement.

University professors did not, however, express the same feeling of togetherness. Commitment to the team was inhibited due to teaching responsibilities at the university at the same time. A stronger sense of interconnectedness was desired and the time required to establish it was unavailable.

Teachers' interviews recognized and appreciated teamwork among the university students and between students and children. As a result of insuffient opportunity to collaborate with professors, some teachers were unclear about program expectations and so requested more communication prior to the onset of the following year's program.

Responsibility. An examination of this theme occurred in three contexts: (a) student teacher responsibility for pupils, (b) professor responsibility for student teachers, and (c) teachers' responsibility for student teachers. Journals also indicated that students came to understand a fourth context of responsibility related to working with their peers. Working with children from five to 12 years of age made students aware of values inherent in some activities, and the problems and demands relating to their combined responsibility for organization. "Smaller children watched the bigger children so it was a good idea to have mixed ages." "We scratched our dance as set out and tried a new version. We used a full circle and then it didn't matter how many were in the group." "Children generally know their limits and set their own parameters."

Professors were reminded of their responsibility to provide students with essential information needed before entrance to the school setting. An even split of responsibility was suggested to allow for this input during the on-campus university coursework and to interweave it into the in-school experience from the onset of courses. Teachers' responsibility for the presentations of practical applications in exchange visits and joint lectures in the university setting or in the school were suggested to enrich and dovetail student experiences. The feasibility of implementing these suggestions rests on shared commitment and responsible attitudes in staff members at university and in schools. Shared responsibility for the integration of subjects and presentation of lectures has not been extensively developed. One teacher expressed the need for all school staff to be present at planning meetings and to feel some responsibility for evaluation from the school's perspective. In their turn, the teacher educators ask, "Are we able to respond in our own institution to ideas of integration and to combining teaching efforts in a responsible way or can such ideas evolve more readily in a university-school atmosphere?"

Resource sharing. Undergraduates perceived most teachers as models, as fountains of resources and ideas, and as mentors who would promote realiza-

tion of their hopes and aspirations. "This teacher has so much knowledge and ideas for us to learn. She said 'Just remember you and I are the same, just I've been at this longer!' Her positive attitude is contagious." "The teacher's attitude was extremely positive not only about us but about the class in general. She was so cooperative and made us feel we could do as much or little in teaching as we felt confident to do." The need for practical ideas was clear in student journals. They appreciated the generous sharing of teachers, much as Fuller and Bown (1975) predicted for this survival stage of teacher development.

Problems and hindrances to collaboration. Sharing, making commitments, and developing co-ownership among project participants exposed numerous hurdles. The school team of principal and teachers established itself throughout the fall term. By the end of the term university staff members were working in the school team and had time to develop understanding of goals and individual roles in the project. Despite this preliminary work together and the fact that staff members participated in the program by choice, one student noted: "If the school experience is to be continued, the teachers involved should be sure they want students in their classrooms. Lack of communication and cooperation made this all so *frustrating* and difficult!"

The collaboration of school and university staff allowed a professional exchange that was stimulating to those involved, but students in one module perceived a need for clarification of expectations. "The teacher and professor should have met together to discuss what was expected from the learning centres." Professors and teachers need to coordinate the content of university module courses with school agendas, and to synchronize perspectives and expectations. Journal entries comment on the "lack of correspondence between the professors and the teachers in schools." The scheduling of assignments within the module was a concern raised by both students and teachers. "Since centres were set up late in the program there was little time to observe the centres in action."

Collaboration and subject integration in the program failed to meet certain student expectations. "In spite of the attempts of those who set up this course, I felt the various subject areas were still quite fragmented from each other." "All three courses should tie the worlds of the school and university together with all professors listening to our experiences and our grievances." Clearly, it is easier to choose participation in collaboration than to commit adequate time and to give priority to it.

Benefits. The reflections of the students revealed moments of illumination, confidence, and pride in themselves and their abilities to meet children's needs and to begin to teach. The shared experience with school staff and children often made them feel a sense of belonging that they also sensed in the children. "The children began to encourage one another. A team spirit really developed."

To take further advantage of the rapport that had been established, one student stated, "I would like to work with the teacher for my third year teaching. The first week would not have to be spent getting to know a new teacher and a new school. This valuable time could be spent getting to know

the children. Her teaching methods are ones that I would like to incorporate into my life." The fruitful collaboration had begun and future time together was desirable.

Interaction

Discussion in this section focuses on interactions among individuals who participated in the collaborative in-school program. The observed effects of those interactions are indicated in student journals and interviews that were conducted with undergraduate students, school staff teams, and professors.

Phases of student socialization. Observation and dialogue, documented by university students in their journals, yielded certain insights regarding socialization into the role of the teacher in its initial phase. "We went to the school today, my partner and I, to get a preview of the environment we will be exposed to and to meet the teachers we will be working with. I'm extremely nervous. Will I ever make it as a teacher?"

Anticipation was followed by a concern about the reception by the resident teacher and students. A warm welcome appeared to ease initial anxiety considerably. "The teacher did everything she could to put us at ease. This made us relax somewhat."

What appeared in this study as a third socialization phase was the students' growth of confidence regarding ability to work effectively with children to promote learning. Teachers facilitated this growth of confidence by providing opportunities to work with individual and small groups of children on learning assignments. Students observed that "I circulated around the class. One child was especially struggling and I helped him and encouraged him through the problems ... the method was effective because after the first few problems, he was able to answer the questions on his own.... This is very encouraging."

Interactions with small groups of children helped to prepare university students for a fourth phase that involved teaching the class as a whole; however, nerves hit again according to student reports. "Even though this isn't student teaching, I'm really nervous about teaching the class." Team teaching involving two or more university students in the planning and presentation of lessons seemed to ease the anxiety for many. "It's great to work with a partner in presenting a lesson; if one person forgets certain details or students' names, the other person pitches in."

An evaluational phase followed. Students themselves intuited what worked instructionally and what did not. "I saw many things in our presentation that were effective and other things we could have improved upon." A number of teachers were helpful to students in assisting them to learn how they could improve their presentations not only in hindsight, but also as students planned their presentations: "Our coop teacher gave us some excellent ideas and guidelines for our language centers."

Students valued participation in the collaboration program for the direct and long-term benefits that they perceived. "Seeing the classroom relations as a participant-observer is directly related to all my other courses." "Through planning a lesson, evaluating the learning center and teaching a small group of students we are getting a feel for our third year practicums."

"Because we now have some experience I'm not so afraid of next year's practicum."

In the collaboration context, second-year education students revealed growth in each of Fuller's stages of professional development (Fuller & Bown, 1975).

1. Preoccupation with their own adequacy: "My problem seems to be a lack of confidence mainly because of lack of experience. It all seemed to come naturally once I started teaching the class."
2. Mastery of the teaching task: "Since we know the evaluations from the times before, every lesson becomes smoother and more polished." "I can see what we need to do to improve the lesson, and what events were beyond our control."
3. Impact on pupils: "Children must see a purpose in what they're doing for them to consider it worthwhile." "Our activity was not challenging enough for the students. They were bored and because of this they did not do what they were asked." Continuous journal reflection on the teaching act showed the concern of undergraduates for their personal relationships with children and with the teachers.

University students and children. Students experienced empathy for the children. "Watching some of the children, I can't believe how they diddle around, but I'm sure I was the same!" They now viewed childhood through different lenses as was stated, "The activity gave us a lot of insight into what it's like to be an 8-year-old." Viewing childhood as a prospective teacher, approximately 10 years after their own experiences as an elementary student, was viewing foreign territory.

What appeared to impress students predominantly in their interaction with children were differences among members of the same class, as some stated, "Students ranged from a very high level of scholastic ability and attentiveness to a very low and disinterested attitude. The less interested students did little to participate." "What struck me most were the differences between children not only in their ability level, but in their level of self-confidence."

Students valued the children as people. They were elated when children invited them to join in their activities. "I felt honored. I enjoyed playing the concentration game with them because we got to know one another better. We began to feel friendly and comfortable together." Many novices in our program revealed that, like Lortie's teacher (Lortie, 1975), their essential reward for effort exerted came from the children. "All the work paid off when I saw how eager the children were to participate in the learning center." Some wrote of sharing the joy of work well done with the younger students. "It felt good to work on an activity that was our own creation—it made me feel worthwhile. The activity was enjoyable for the students and us." "It's great to see bonds forming between the students and the kids." In almost all cases university students established rapport with children in classrooms early in the program.

Peer interaction. Among participating university students, peer interaction enhanced program benefits for the majority of students. They experienced failure together. "We all spent too much time verbally giving instruc-

tions." They derived support and a sense of belonging from pairing. "Choosing a partner brought a sense of belonging in a group of total strangers."

Some student liaisons experienced stress due to lack of interpersonal sensitivity. "My partner monopolized time with the children and teacher. When it comes to preparation she's very irresponsible—I have to do all the work." Wounding and resentment developed in such circumstances. Some student groups found that it was not an easy job to work effectively together. The project placed them in interactive situations demanding flexibility and adaptability that was difficult for some students. "We had problems making up the lesson plan as my partner and I had many different ideas and could not incorporate them all. I found it difficult to teach the lesson as I found there was not enough cooperation."

Problems/hindrances to interaction. Concerns centered on the lack of adequate preparation of novices. These students were not adequately equipped to analyze their interactions and observations of student and teacher behaviors. Their need for information regarding peer socialization in classrooms, effects of moral development, early development of self-concept, and the influence of self-conceptions on learning was evident in journal discussions of observational experiences. University students and teachers observed that instruction in lesson planning and presentation was also necessary prior to students' participation in the in-school program.

Where negative evaluations were reported, the positive effect of interaction was also observed in the reflection it stimulated. The program was designed to further the growth of insight concerning what types of interaction would benefit both schools and the university. Participants reflected on their own involvement and that of others as they evaluated the effects of the program.

Reflection

From the very first planning of the collaborative school-based teacher education program, a key principal was that "Reflection is essential to growth." In lay terms, reflection means meditation, contemplation, mental consideration, stemming from the Latin root *reflectere*, re + flectere (to bend). The rebending through mental consideration suggests that we consciously reshape ideas and experiences, reconstructing and rebending in our minds the received ideas to explore new possibilities, new approaches to problems. Schön, in *Educating the Reflective Practitioner* (1987), seems to have this recasting meaning in mind as he discusses learning a profession by practicing it while "reflecting-in-action." Schön suggests that professional education needs to place neophytes in field work, assuming that they encounter the unexpected and reflect on it. "Reflection-in-action has a critical function, questioning the assumptional structure of knowing-in-action. We think critically about the thinking that got us into this fix or opportunity; and we may, in the process, restructure ... ways of framing problems" (p. 28).

This need for students to reflect-in-action demanded "a willingness to critically reflect on 'What is?' and to consider 'What could be?'" and was explicitly stated in the program goals agreed to by principals and professors at a meeting November 5, 1987. To promote reflection, students were asked

to write in journals after each school visit. An examination of journal entries and later interviews sheds some light on how students were stimulated to reflect-in-action by experiences with the teacher-professor-teaching experience triad.

Observation. One of the triggers to reflection was the opportunity to be in a school and regularly observe an experienced teacher working with children. Even the initial view of the school structure contrasted with earlier school experiences to stimulate reflection on the impact of the physical environment on children. "Arrived at the school and was immediately impressed with the openness and inviting atmosphere of the structure." The schools involved are designed to provide numerous nooks and crannies, inviting teachers to break whole class groups up into smaller groups, challenging children to accept the responsibility of independence. This was a new experience for most students, and forced reframing of traditional conceptions of using space to change teacher-student relationships.

This reflection was extended by observation in classrooms as students saw new ways of working with children. "I was amazed at how fast the pace of the afternoon was. 'Lessons' seemed to be so short compared to when I was a kid. I expected 'lessons' to be one half hour long, but each activity was only 10-20 minutes. There was a lot of movement in the classroom." What was commonplace for the experienced teacher was a surprise to the observant novice and forced some reframing of conceptions of children.

Experiencing teaching. Although observation provided students the opportunity and perhaps the leisure to reflect, it was the stimulus of experiencing teaching that resulted in the most journal reflection. As students encountered problems, they needed to reflect-in-action: "How can I get the kids to slow down and absorb the material rather than passively reading? I want them to think about and analyze or reflect on what they have just done." "Our main problem is lack of control and getting and keeping children's attention when we are giving directions." Schön (1987) seems to be correct in expecting that problems in teaching practice would stimulate student teachers' reflection.

Teachers' roles in reflection. Although students' journal entries indicate that they can reflect on their experience at home alone, there were also frequent comments on reflection in conversation with their teachers or as a result of things teachers had said or done. Sometimes students had developed clear goals of their own and found that they did not agree with those of the teacher. "I am now beginning to see my future goals and approaches to teaching in comparison with those of the teacher and find that they clash in some respects," or again, "I find it hard to see any substantial interaction between the teacher and students since her method is quite rigid and structured." More frequently, however, students noted how much they were learning from the teacher as they talked over problems and ideas with them—closer to Schön's reflection-in-action.

Professors' roles in reflection. For the combination of experience and reflection to occur, students need some conceptual tools to use to analyze experiences and some value commitments to use to evaluate their teaching. Students used some of the concepts introduced in courses to reflect-in-action.

"I really became aware of the hidden curriculum—to decide what type of teacher I am, what goals are at the top of my list, and what methodologies I favor." However, these goals and ideas now had practical value so important at the beginning stage of becoming a teacher, as Fuller and Bown (1975) have noted. One student pointed out the importance of the teaching experience in the preservice courses. "The difference is feeling you're in the program to study versus you're learning to be a teacher." Further, some students realized that how they taught was affected by specific concepts from courses. "Our study of social action, reflection, etc., is good in that it continues to influence our teaching unconsciously as well as directly." Critical reflection assumes evaluation, the application of values and beliefs. Some courses seemed to have provided this to the undergraduates.

Factors inhibiting reflection. The lack of rapport appeared as an important factor that seemed to inhibit reflection. On occasion there was failure to develop a relationship in which students, teachers, and professors felt safe to risk sharing concerns, doubts, feelings of uncertainty and inadequacy, or anger. It appeared that a lack of desire kept participants at a distance from one another. "The only time our teacher has appeared pleased occurred today when we indicated we will not be at the school next (Reading) week."

Time was frequently lacking as students sought to discuss with teachers their observations and experiences in the schools. Teachers also pointed out that some students arrived after the children, so there was no time to plan together. Teachers and students often had commitments after school that cut short the time needed to reflect together on students' teaching and problems or questions. Similarly, professors were perceived not to be in the schools enough where teachers and students could reflect-in-action with them about assignments, problems needing clarification and solution, and differing points of view. It seemed that the module was a small part of the professors' assigned teaching load. Consequently, they were unable to be in the schools to interact with students and teachers during the richest time for reflection on the events and their teaching.

All participants experienced the need for ongoing dialogue for the purposes of planning, evaluating, providing feedback, and facilitating reflection. Teachers observed that "keen, eager prospective teachers have questions and time for dialogue must be scheduled." "There was a lack of time to communicate with students prior to and after visits." Students regretted the "lack of opportunity to discuss the lesson with the instructor prior to that day." One student commented, "The teacher said she was sorry she doesn't have more time for us. We said we realize she is very busy; we do not want to impose on her." Time for conferencing needs to be provided if the full value of reflection is to be achieved. Action is required to access financial assistance; collaborative funding by the university, school board, teachers' association, and government should be investigated.

Summary. These excerpts from student journals and interviews indicate that some reflection-in-action was generated by the concurrent school experience in being a teacher, secure in having the support of an experienced teacher, while taking courses that provided concepts and values aiding reconsideration of these experiences. Factors that made a positive contribution to

reflection seemed to include: (a) direct experience in teaching and the problems that result; (b) opportunity to observe a teacher with a particular teaching style, sometimes at odds with that of the student; (c) time to sit down in a relaxed, trusting setting and talk over with their teacher or professor the problems students encountered in their teaching; (d) encountering differences in teaching philosophies among the student, teacher, professor triad; (e) writing in their journals about that day's school experience, and reading the feedback received; (f) learning concepts and values in courses that could be used in analyzing and evaluating teaching methods used by students and teachers; (g) having assignments that required designing plans for achieving educational goals, then trying out those plans and discussing and evaluating results with teacher, peers, and professors; (h) providing manageable (weekly) increments in student responsibility, enough to stimulate reflection on problems encountered, but not so much that the complexity of the task overwhelmed analysis, reframing, or reflection-in-action; (i) placing a pair or small group of students with teachers so there was another beginner with similar course background and assignments with whom to reflect on their school experiences; (j) providing assignments that involved students in cooperative planning and teaching to foster discussion, debate, comparison of interpretations of the situation, and nonthreatening feedback on their teaching; (k) freeing teachers from an evaluative role so that students could openly discuss with them problems, failures, uncertainties and anxieties encountered in their initial teaching experiences; (l) encountering a wide variety of different kinds of children with varying abilities, needs, home backgrounds, school histories, and attitudes in both special and regular classrooms so that students were pressed to reflect on purposes of schooling and varied methods appropriate for children's differing characteristics; and (m) taking a set of common courses with common assignments and school experiences helped develop a sense of community in which students were more able to risk sharing their problems, dissenting views, worries, and emerging ideas. Where teachers and professors were included in this community, wider sharing sometimes heightened reflection and insight.

Implications and Further Questions

Collaboration, Commitment, and Interaction

As Montagu (1966) has indicated, "The critical social and educational problem of to-day is one of learning how shared relationships may be fostered and freedom of inquiry accelerated. It seems to me, first, that this must be done through the schools, and second that there must be virtually complete change in our attitudes towards education" (p. 118). The collaborative community of educators involved in this study raised questions about how time can be made available so that roles of professors and teachers can be redefined. How can teachers assume more of a professorial role and professors more of a teacher role while student teachers have opportunity to become reflective teachers? The initial planning process included professors, principals, and teachers. After four months of planning, students appeared in university classes and learned they were to spend a half-day or so each week in a classroom in addition to attending the lectures that they had expected would constitute

the program in its entirety. They were now expected to arrange transport to and from schools, keep journals, plan lessons—all for no additional credit. There was considerable feeling of injustice concerning this requirement, but little surprise that they had had nothing to say about these plans for use of their time. In spite of the inconsiderate treatment that they had been subjected to, students demonstrated their ability to suggest good ideas as the term progressed, and during the final evaluations showed concern for others in addition to themselves in discussing directions for program change. The paternalism of past university practices in precluding student participation in program planning has been costly. Much stands to be gained from inviting student contributions to planning.

Willingness to cooperate with others in the context of such a program must be stated as a requirement for participation. Planning sessions involving professors and teachers need to be conducted prior to the onset of in-school sessions so that expectations for student performance can be mutually delineated. Professors need to assume responsibility interdepartmentally for the provision of module course instruction and related in- school learning. Student peer group meetings and school staff meetings need to be scheduled to establish support systems. School administrators need to stimulate reflection on developments in all participants.

The sense of belonging that arose from shared experience with the school community, staff, and children was captured by some undergraduates. They reflected on moments of illumination that revealed their ability to meet children's needs and to begin to teach, as well as awakening confidence and pride in themselves. The desire to follow this term course experience with their full-time student teaching practicum was seen as advantageous and a saving of time for both participants. In fact, several students were able to arrange their return to the same classroom for their practicum.

Reflection

An important program goal was to stimulate student reflection on teaching. Data from student journals, student interviews, teacher evaluations, and professors' interviews provide evidence that there was considerable reflection during the program. Particularly significant was the concurrent experience of taking courses while teaching children. Where student teachers in other practica have no ongoing courses, they often seem to focus solely on learning to teach by modeling their cooperating teacher. Although they have previously taken courses that have equipped them with concepts to use in planning and teaching, students largely take over their teacher's existing program, classroom management procedures, and materials. The enormity of the task and its exhausting demands plus the evaluative role of their teacher frequently lead to student imitation of their teacher in order to survive. This reduces the opportunity and inclination to critically reflect. However, because students were only spending a half-day each week in classrooms, and not teaching for all of that time; because their teacher was situated in a nonevaluative role; because one or more peers were sharing the responsibility and the teaching; because professors were discussing their experiences and course concepts with them both in the schools and in class;

and finally because some course assignments were done in classrooms, the concurrent "study plus teaching" program seemed to stimulate and support more reflection and openness about teaching.

In some instances, students frequently encountered differences in values and beliefs about teaching among their teachers and their professors. This seemed to generate cognitive dissonance, a need to seek resolution of competing ideas, to reconstruct their beliefs to accommodate these varying ideas. Sometimes journal entries indicated growing skepticism about the feasibility of a professor's pet idea. At other times students rejected teachers' handling of discipline. Often they left the issue open, uncertain of how to resolve it, but allowing it to remain a topic for further discussion, study, reflection, and experience.

The impact of time on reflection was a continuing concern to participants. The lack of time was felt as teachers and professors attempted to reflect and to discuss the program and students' experiences. Because teachers had no time to participate in the courses, nor to question professors about ideas students were bringing to their classrooms, they were less able to use course concepts to reflect on their own teaching ideas and to share in student use of those concepts in their reflections. To be truly collaborative and achieve inservice growth, such a program needs to assure adequate time for teacher participation.

Finally, rapport among students, teachers, and professors often affected reflection. When students and teachers developed a warm relationship, more open sharing of questions about teaching took place. Similarly, where students in a classroom became good friends, they provided emotional support for each other's thinking as they jointly encountered problems in their teaching and more critical reflection was likely to occur. Students who perceived their relationships with professors to be open and encouraging of exploration of divergent ideas reported they were more likely to use class time to go beyond learning concepts, and to use concepts for analysis and evaluation of school experiences. Time seemed to be a major contributing factor. When other demands pressed on students and teachers, good rapport seemed more difficult to develop, and the openness needed for critical reflection was less likely to occur.

Conclusions

The first section of this report traces the history of the collaborative school-based teacher education project back to plans by members of the Department of Elementary Education to work with school staff to reconstruct the teacher education program. Much of the foregoing evaluation report focuses on details of the four-month school-linked modules project. However, it is important not to lose sight of the broader purpose of renewal of the whole four-year BEd program.

It seems clear from the conclusions reported above that some structural aspects of teacher education programs stimulate and some inhibit students' reflection-in-action. It seems desirable to plan concurrent practica and coursework. This idea is further supported by research reported by Tickle (1987) and by Zeichner and Tabachnick (1985). Tickle's students at East

Anglia University are attached to schools one day a week during all four years of their program, and Tickle describes the projects students develop that enable them to go beyond learning the nuts and bolts of teaching to research on teaching as part of a teacher-student-tutor triad. The conception of the school attachment as an opportunity for research and reflection in which apprentice, master teacher, and tutor-researcher collaborate in systematic study of teaching changes the orientation of students from learning to teach the way a particular experienced teacher teaches, to studying how children respond to different approaches to teaching.

At Wisconsin, Zeichner and Tabachnick use a course in the sociology of education to prepare students to study a school staff's norms, beliefs, and practices and the history of their development as a way of reorienting students to a heightened awareness of the tacit beliefs and usually unquestioned norms and practices that prevail. The conceptual tools of sociology enable more analytical and evaluative reflection on proffered teaching norms shared by the teaching staff.

The results of these two programs seem consistent with the present study and imply the need to structure teacher education programs to provide concurrent practica and coursework with explicit activities planned to promote reflection, such as journal writing, assignments calling for evaluation of students' teaching, research on specific teaching approaches, participation in research teams, and the like. The risks of failure are heightened by the disempowerment of students when experienced teachers and professors are seen to be sitting in judgment on early efforts to make ideas work in practice, and to reflect on those efforts. A nonevaluative practicum offers more support for reflection than one where student success is rated by teachers and professors.

It is noted in the section on Reflection that students were helped to reflect-in-action during their school visits by using some of the concepts introduced in their concurrent education courses. However, students clearly were limited in their reflections by not being able to more effectively apply concepts from the sociology of education to explore the ideological and the sociopolitical positions that teaching beliefs and practices represented. Hence much of the hidden curriculum remained hidden. Nor were students adequately conversant with alternative philosophies of education that depict different roles for schools and different supporting belief systems. This prevented students from engaging in substantive reflection on their evolving personal philosophies and critiquing personal practical knowledge regarding teaching. Similarly, their limited grasp of principles of child development and learning theory restricted their ability to question the appropriateness of methods, materials, and content being used in the schools.

It seemed that much of the knowledge developed in these fields had been too abstract, too removed from the undergraduates' own experiences, and from children and teachers in classroom groups for it to be available as a conceptual tool useful in analysis and critical reflection on their school experiences. One of our failings has been the omission of staff from Educational Foundations and Educational Psychology in the planning and undertaking of the project. The consequence is too narrow a reconceptionalization

of the BEd program. Too heavy a focus on the marriage of school and campus in the module curriculum and instruction program has indicated a narrowness in this project.

With expressed administrative support and funding, staff inservice commitments for one year or for the full four years of the BEd program could be focused on the collaborative teacher education plan. Inservice pressures can lead to overextension of staff energies. Protection may be found in facilitating the commitment of university professors and school staff to a focused professional collaboration. Shared research roles and cooperative teaching experiences could stimulate all staff members involved.

Note

We wish to recognize the important contribution made to this study by our research assistant Mary Oliver. Her analysis and interpretation of the Interaction theme and her participation in the research group were most valuable.

References

Blakey, J., Chamberlin, C., Everett-Turner, L., Halabisky, G., Kysela, G., Maaskant, F., Massey, D., Massing, C., McNay, M., Mochoruk, M., Sande, D., Scott, N., Stephenson, E., & Tucker, D. (1989). *Sources of teachers' perspective and decisions: Implications for preservice and inservice education*. Edmonton, AB: Alberta Education.

Craig, T. (1984). *Towards an understanding of the life-world of the first year drama teacher*. Unpublished doctoral dissertation. University of Alberta.

Feiman-Nemser, S., & Floden, R.E. (1986). The culture of teaching. In M.C. Witrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 505-526). New York: Macmillan.

Fuller, F., & Bown, O. (1975). Becoming a teacher. In K. Ryan (Ed.), *Teacher education: The 74th yearbook of the NSSE* (pp. 25-52). Chicago: University of Chicago Press.

Iannacconne, L. (1963). Student teaching: A transitional stage in the making of a teacher. *Theory into Practice, 11*, 73-80.

Lacey, G. (1977). *The socialization of teachers*. London: Methuen.

Lortie, D. (1975). *Schoolteacher*. Chicago: University of Chicago Press.

Miklos, E., Green, M., & Conklin, R. (1987). *Teachers' evaluation of their preparation for teaching*. Edmonton, AB: Alberta Education.

Montagu, A. (1966). *On being human*. Toronto, ON: Prentice-Hall.

Newberry, J.M. (1979). The beginning teachers' search for assistance from colleagues. *Canadian Journal of Education, 4*, 1, 17-27.

Schön, D.A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the profession*. San Francisco, CA: Jossey-Bass.

Tabachnick, B.R. (1980). Intern-teacher roles: Illusion, disillusion, and reality. *Journal of Education, 162*(1), 122-137.

Tabachnick, B.R., & Zeichner, K. (1985). *Individual and contextual influences on the relationship between teacher beliefs and classroom behaviors: Case studies of two beginning teachers in the United States*. Paper presented at ISATT 1985 Conference, Tilbury University, The Netherlands.

Tickle, L. (1987). *Learning teaching, teaching teaching*. London: Falmer.

Zeichner, K.M., & Tabachnick, B.R. (1981). Are the effects of university teacher education washed out by school experience? *Journal of Teacher Education, 32*(3), 7-11.

Zeichner, K., & Tabachnick, B. (1985). The development of teacher perspectives: Social strategies and institutional control in the socialization of beginning teachers. *Journal of Education for Teaching, 11*(1), 1-25.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 2, June 1991, 157-166*

COLLA JEAN MacDONALD
University of Ottawa

Elementary School Teachers Explain Why They Do Not Use Creative Dance in Their Classrooms

Creative dance programs in elementary schools have been found to benefit children's development and to provide an enjoyable and expressive activity for children (MacDonald, 1989). Such programs have also been shown to act as an innovative method of teaching and integrating the elementary curriculum. Despite these positive contributions, little if any creative dance is now being taught at the elementary school level (Courtney & Park, 1981; Fowler, 1989). This article presents findings regarding why elementary school teachers are not currently implementing this art form in their classrooms. Elementary school teachers were recruited and selected for an interview, and the data gathered in this interview were used to identify teachers' reported attitudes and practices about teaching creative dance. Teachers' reasons for not implementing creative dance formed three main categories: attitudes and beliefs, knowledge or skills, and resources or facilities. The findings regarding attitudes and practices indicate that support and inservice that address teachers' needs in the areas of attitudes and beliefs, knowledge or skills, and resources or facilities are necessary. The findings also suggest that providing intervention addressing teachers' expressed needs could go a long way toward ensuring that more children are exposed to this noncompetitive, nonthreatening activity.

Introduction

Have you ever seen children so totally involved in what they are doing that their heads are wet with perspiration and their faces beam proudly as their bodies interpret and express a life situation? Have you observed schoolchildren moving to music and using their bodies to solve a mathematical equation? Have you watched children get so excited while manipulating a dance prop that they do not want to stop for recess?

These are just some of the experiences of elementary school teachers who have implemented creative dance in their classroom. As is the case with the arts in general, though, little if any creative dance is now taught in the school system (Courtney & Park, 1981; Fowler, 1989). This article explores why, despite these and other positive experiences, elementary school teachers are not currently implementing this art form more widely in their classrooms. It is first necessary, however, to define creative dance and discuss its importance in the curriculum.

Colla MacDonald is an assistant professor and program leader for the Primary-Junior Teaching Division of the Faculty of Education. Her research interests include the use of the arts as an innovative method of teaching and integrating the curriculum, educational change, and student teachers' relationships with associate teachers and professors.

What Is Creative Dance?

In this article, *creative dance* refers to activities expressed through the medium of the body that focus on inner thoughts and feelings and enhance the expression of those thoughts and feelings. Creative dance emphasizes creativity, problem solving, and the expression of thoughts and feelings; it also involves participants physically, emotionally, and intellectually. For example, mathematical fractions might be taught by rhythmic chanting and movement, poems interpreted through creative dance, or the meaning of the planets and their position in the universe instilled by creating dances.

Creative dance uses basic dance techniques borrowed from other dance forms, including modern dance. Creative dance is a nonthreatening, noncompetitive activity that recognizes and encourages individual differences while challenging participants to improve physical skills and aesthetic expression. Creative dance is easily taught in gymnasiums, multipurpose rooms, or classrooms, and it requires no previous dance background of students or teachers.

Why Is Creative Dance Important?

I have studied and taught creative dance in elementary schools and universities in different Canadian cities and provinces for 12 years. As a result of this experience, I believe that creative dance can play a vital role in elementary schools both as a focus within the arts curriculum and as a vehicle for learning. I also believe that creative dance can be used as an innovative method of teaching and integrating the elementary curriculum.

There is considerable evidence that creative dance is of extensive value and importance to children (Best, 1985; Boorman, 1973; Courtney, 1980; D'Houbler, 1957; Laban, 1968; MacDonald, 1989; Riley, 1984). First and foremost, children experience feelings when they participate in this art form. Moreover, these feelings are particular to and inseparable from the art form in which they are expressed (Best, 1985). The feelings creative dance gives children, therefore, cannot be experienced in any other way, and only those children who have experienced creative dance can appreciate its physical, intellectual, and emotional impact. As a result of this total involvement, a dialectic is established between the child's inner and outer worlds. Physical movements are external representations of what is occurring internally, and because these movements could not take place without inner experience, the inner experience cannot be denied (Best, 1985).

Creative dance programs in elementary schools have been found to benefit children's development and to provide an enjoyable and expressive activity for children (MacDonald, 1989). The entire process emphasizes exploration, problem solving, and the interpretation of life issues with the intention of enhancing the development of the whole child: the child's inner world, outer world, and aesthetic expression. Encouraging the expression of individual differences, creative dance also gives children an opportunity to explore their views on life issues, on the entire human condition, and on their own human condition. This exploration may in turn deepen their understanding of life's experiences and increase their sensitivity to these experiences (Best, 1985).

Creative dance can also facilitate children's personal development in significant ways. For example, through this art form, children can discover that many situations have a number of meanings, experiment with a variety of solutions to problems, and learn to accept and reject one another's ideas (Boorman, 1973; D'Houbler, 1957; Riley, 1984). Creative dance gives children an opportunity to solve problems by exploring situations, testing hypotheses, and discovering solutions. Moreover, creative dance gives children myriad opportunities to communicate, develop social skills, and increase their self-confidence.

Creative dance may also provide ways of learning that could not otherwise occur (Boorman, 1973; Gardner, 1985). As a result, it is possible to integrate creative dance into the teaching of various core subjects in the curriculum. Put more strongly, in a curriculum espousing holistic concepts, creativity would be of paramount importance in all subjects, the arts would be integrated into all core subjects, and creative dance would have an integral place in the curriculum (Boorman, 1973; D'Houbler, 1957; Laban, 1968; Riley, 1984).

More precisely, creative dance would be taught not only as a subject but also as a method of learning. Like learning centers, this use of creative dance would integrate subjects across the curriculum; in addition, it would develop the whole child. Miller (1988) endorses the assumption underlying this article, namely, that creative dance is more than a physical process, that it involves other aspects of our being. In addition, Miller supports Gardner's view (1985) that denying children's physical development can negatively affect their full development—including intellectual development.

In sum, creative dance can help children reach their full potential, for it encourages the development of the whole child by enhancing creative exploration and facilitating emotional expression. Moreover, if children feel good about their bodies and their movements and if these outward expressions help children realize and interpret their inner feelings through the medium of the body, everything they do will be affected.

Research Design

Elementary school teachers in a mid-sized Ontario city were recruited and selected for an interview. Recruitment took place in university courses, practice teaching sessions, and professional development workshops in which I was involved. In these various environments, I presented information about my research, described its objectives, and explained the criteria for participation:

1. Participants had to currently be teaching Jr. Kindergarten, Kindergarten, or Grades 1, 2, or 3.
2. They had to have taught Jr. Kindergarten, Kindergarten, or Grades 1, 2, or 3 for at least one year.
3. They could not currently be teaching creative dance as the term is defined in this article.

Thirty-three teachers expressed an interest in participating. Of these, 19 women and one man met all of the criteria. All participating teachers were asked to take part in an audiotaped interview. Because of its convenience, the setting for most interviews was the classroom. Interviews ranged from 20 to

30 minutes in length. The interview questions focused on participants' teaching and dance background, their attitudes toward and practices regarding creative dance, and their intentions, beliefs, anxieties, and needs with regard to the teaching of creative dance. Questions were posed in such a way as to both increase my understanding and improve my interpretation of teachers' responses and make me more aware of teachers' needs regarding the teaching of creative dance. In the interview, teachers spoke freely about their opinions, problems, anxieties, beliefs, needs, and frustrations concerning creative dance.

This audiotaped interview was transcribed in full. I read, analyzed, and reflected and deliberated on the transcripts' contents in an effort to identify and determine acceptable and meaningful terms that would enable me to understand, organize, and explain findings. Teachers' reported attitudes and practices were documented using their own responses, and their needs with regard to teaching creative dance were identified and interpreted.

I next met with each teacher individually and provided him or her with a copy of the transcript. My purpose at this time was to get each individual's reactions and to record any necessary additions, corrections, explanations, or clarifications. With the exception of minor grammatical concerns, no major problems arose, and these issues were immediately addressed and wording changed to clarify the teacher's intentions. The content of the individual transcripts was categorized and interpreted.

Results

This section presents information on teachers' background characteristics and explanations for attitudes and practices related to both creative dance and the teaching of creative dance.

Background Characteristics

Of the 20 teachers interviewed, six taught Jr. Kindergarten, five taught Kindergarten, six taught Grade 1, eight taught Grade 2, and three taught Grade 3. Eight of the 20 taught split grades, as a result of which the total number of grades taught was 28, not 20. Teaching experience ranged from four to 27 years, with an average of 13.4 years of full-time teaching.

Eleven teachers had no previous training or experience of any kind in dance. Nine had some exposure to dance; however, for every teacher but one, this training or experience was very limited. For example, one teacher had taken a noncredit social dance course at a community college; one had taken six jazz dance lessons, a few had taken some creative dance as part of a Ministry of Education course, and another had been to a professional development workshop on creative dance.

All teachers reported they had never taught creative dance as part of their classroom program. Nevertheless, all also said they believed creative dance should be included in the elementary school curriculum. When asked why they felt this way, all expressed the belief that creative dance was important to the total development of the child. For example:

> Teacher 4: I think it is important, especially because some children who aren't good academically might excel in that area ... but children that aren't really good at other things can really come out of their shell.

Teacher 6: Well ... because creative dance allows the child to express ... ideas and ... opinions not only mentally and verbally but through movement.

Explanations for Attitudes and Practices

Given that not implementing creative dance was one criterion for selection in the study, the unanimous belief that creative dance should be part of the curriculum was especially interesting. Teachers' reasons for not making creative dance part of their program thus took on added significance. Three categories of explanation emerged: attitudes or beliefs, knowledge or skill, and resources or facilities. In each category, responses could be further differentiated into three to five sets, each set representing a particular feature or dimension of the category. For example, in the category of attitudes or beliefs, a set of responses all expressing lack of confidence emerged. It was necessary for an explanation to be mentioned by at least two teachers to qualify as a category.

Attitudes or beliefs. This category addressed teachers' personal attitudes or beliefs regarding the teaching of creative dance. Four sets of attitudes or beliefs were identified: *confidence*, *priority*, *time*, and *apathy*.

Confidence referred to responses indicating that teachers believed that they did not have enough confidence to implement creative dance. Eleven teachers cited this type of belief as an obstacle. For instance:

Teacher 16: Yeah ... I think there is an element of [discomfort] (laughs) when you haven't done a lot of it yourself.

Priority responses included those indicating a belief that creative dance was not important, for example, because the Ministry had not emphasized it in the curriculum or it was not important relative to the core subjects. Eleven teachers cited this attitude as an obstacle to implementing creative dance; responses included the following:

Teacher 6: There have been a lot of these documents out lately, and every time a document is out, my Board goes around and does a lot of mini-workshops on the document, making sure that it is implemented, making sure we understand what is going on. So in the last few years, we have had workshops on art, we've had workshops on drama, we've had workshops on shared discovery. Nothing ... if there is nothing on creative dance, then teachers take the priorities. The priorities are the Ministry priorities.

Teacher 11: It's never seemed to have been a very high priority with anyone with the Board.

The third set, time, included responses indicating that lack of time was one of the reasons teachers had not taught creative dance; lack of time refers to not having enough time to learn material regarded as necessary, not enough time to look for teaching resources, not enough time to fit dance classes into programs, and the like. Nine teachers cited lack of time as an obstacle to teaching creative dance. The following responses were included in this type of attitude or belief:

Teacher 11: Well ... it's the usual, you know, you just don't seem to have enough time in the day... . And if you say, "Gee, I have to get this done and this done" and there are all these pressures from above, from the principal

or whatever, and this has to be done, and the children have to be able to do this ...

Teacher 19: It's not just the initiative; it's the time, you know, when you're planning this and you're running off that and you are trying to see children for something else and you're writing reports and you're putting up bulletin boards, you know, it's just ongoing all the time.

Apathy responses included those indicating that teachers were not implementing creative dance because they couldn't be bothered to teach it. When teachers were asked why they were not implementing creative dance, only two cited this factor as an obstacle:

Teacher 16: Well, no, I am just thinking like ... in our school particularly ... one of the reasons that was brought up one time in the staff room was that you have to carry [so much] around.

Knowledge or skill. The category knowledge or skill addressed teachers' knowledge or skill as it related to the teaching of creative dance. Four sets of knowledge or skill responses emerged in the analysis: *content*, *methods and integration*, *music*, and *rationale*.

Content referred to responses indicating that teachers identified concern with content—with what to teach—as a reason for not including creative dance in their programs. All teachers identified content as a factor. Following are two content responses:

Teacher 6: I didn't know *what* to do ... I have a group of 20 children; what specifically do I do with those children in my classroom?

Teacher 13: I don't have the background for that at all ... the content, the nitty-gritty of what to do and then where to use it and how to use it.

Methods and integration included responses indicating that teachers had not included creative dance because they did not know how to implement or integrate it. All teachers cited methods and integration as a problem. Here are two such responses:

Teacher 6: I did not know *how* I could incorporate it into my class without taking a lot of time away from other subjects I am teaching.

Teacher 15: Just lack of knowledge of *how* to implement it basically.

Music, the third set of knowledge or skill responses, referred to responses indicating that teachers did not know what music to choose or where to find appropriate music. When teachers were asked why they had not included creative dance in their program, 12 cited music as a problem. For example:

Teacher 15: I don't know where to start; I don't think that I have the right records in my classroom or at my fingertips.

Teacher 4: Materials and music and stuff like that. Because I find that music is hard to find.

The final set of knowledge or skill responses, rationale, included comments indicating that teachers had not taught creative dance because they did not understand why it should be incorporated into their program. Six teachers cited lack of rationale as an obstacle to implementing creative dance. The following comments highlighted teachers' concerns in this area:

Teacher 15: Ah ... we have a section on creative dance in our core program, but ... sometimes I read the objectives, and I don't understand its relevance to creative dance.

Teacher 6: I need to know *why* I should teach creative dance in my classroom.

Resources or facilities. The third category addressed teachers' concern about whether enough teaching materials, resource materials, resource persons, and facilities for teaching creative dance were in fact available to teachers. Six sets of responses pertaining to resources or facilities emerged: *recognition and support from the Ministry*, *reading material*, *equipment*, *instruction*, *support systems*, and *space*.

Recognition and support from the Ontario Ministry of Education included responses indicating that teachers were not implementing creative dance because it was not included in the curriculum or recognized by the Ministry, for example, by way of a support document or inservice training. Fifteen teachers identified lack of direction and support from the Ministry as an obstacle to implementing creative dance. The following responses indicated the importance of Ministry support:

Teacher 11: Well, dancing is such a ... well, at the moment, it is not part of the curriculum ... so if you really want to make it an important factor ... as I say, the Ministry would have to ... really make it a must and make it part of the curriculum.

Teacher 12: But it is something that tends to be skipped over ... because people tend to go to the area of sports ... but I think if there was a Ministry document, then that might make people see just how important it is.

Reading material referred to responses indicating that teachers had not included creative dance in their program because they did not have the appropriate articles, documents, books, or other literature. When asked why they had not taught creative dance, 15 teachers remarked either that not enough had been written on creative dance or that not enough written material on the subject was available to them. The following responses identified the need for reading material:

Teacher 5: But then it is always nice to have a reference to go to, to say "Hey! These are ideas I could try" ... some, perhaps, skills that could be taken a bit further. So that's possibly one of the downfalls which has discouraged me from teaching it.

Teacher 11: And I haven't seen anything that I can remember that ... would help me, that I could say, "That's right, that is something I could use" ... and to give me an idea of where to start, and what I could do, what is something I could do with four-year-olds.

Equipment, the third set of responses related to resources or facilities, included comments suggesting that teachers were not implementing creative dance because they did not have access to the needed equipment (record players, tape recorders, extension cords, props, videos, costumes, and the like). Eleven teachers indicated that equipment was an obstacle to teaching creative dance. Responses included the following:

Teacher 12: And a ... lack of finances to get stuff once you've, once I've thought of things that I'd like to do.

Instruction responses included those indicating that lack of experience in or exposure to instruction in creative dance, for example, in courses, classes, or workshops, was a reason for not teaching creative dance. Eight teachers indicated that this factor was a problem. Comments associated with instruction included the following:

Teacher 15: I think I speak for a lot of people that people are not comfortable with it because we haven't had that in high school or elementary school.

Teacher 17: I just felt I haven't been trained. Even when I went to teachers college back in 1967, we didn't have anything for that.

Teacher 11: I feel if you want to put some emphasis on creative dance, you can only do that through workshops or during PD days.... You want to give some of your time, but if they really want to make this really beneficial and something that will be successful, give us the time and a chance to really do it.

Support systems responses indicated that teachers had not included creative dance because they did not have access to resource persons or support systems. When teachers were asked why they were not implementing creative dance, four referred to this issue of support. Responses included the following:

Teacher 9: It never seemed to have been a very high priority with anyone, with the Board, whether it be the Phys. Ed. consultant or whoever. I mean, that has just never been a high priority with any of them; so ... nothing really has come from them.

Teacher 4: Because I think some principals probably don't see creative dance as Phys. Ed. or as valuable; so you have to know why you are doing it.

Space, the final set in this category of responses, referred to comments suggesting that space had prevented teachers from teaching creative dance. Three teachers identified space as a problem; responses included the following:

Teacher 12: The time element with the gym ... the gym being assigned for 30 minutes twice a week really limits what you want to do and the amount of time you can spend in that area ... lack of facilities, at least what I would consider a lack of facilities.

Teacher 4: You are limited in your Phys. Ed. two or three blocks a week. And in a classroom with 30 desks and you know, one little ... space for the kids to sit. You have to be able, sometimes it sounds good on paper, but when you get in with 30 kids to try to walk around, it's hard.

Findings: Highlights and Implications

To summarize, many teachers did not feel comfortable with or confident about teaching creative dance. They cited not knowing what creative dance was and how it could benefit their school program as reasons for not including creative dance in their classroom practices. With respect to teaching creative dance, many teachers said that they did not know what to teach, how to teach it, or how to integrate it across the curriculum. In addition, the lack of resources or facilities available also affected teachers' practices regarding the teaching of creative dance. Finally, the priorities established by the

Ministry, Board, and school principals had considerable effect on classroom practices regarding the teaching of creative dance.

These findings have certain implications for interventions designed to encourage elementary school teachers to teach more creative dance. First, whatever form of intervention is chosen, it is important to establish a climate of respect for teachers' uniqueness as individuals and to be flexible in response to teachers' questions and concerns. For example, workshop facilitators should establish a climate in which teachers feel free from tensions, stress, doubt, or embarrassment; encourage teachers to choose their own level of participation; and encourage teachers to be self-directing and to adapt or extend the workshop experience to meet their classroom needs.

Second, since understanding the meaning of a proposed change is a vital step in ensuring that change will occur (Fullan, 1982; Regan & Leithwood, 1974), intervention should enhance teachers' understanding of what creative dance is, why it should take place, and how it can benefit teachers' programs. In workshops, facilitators should explain the role creative dance can play in the elementary school program.

Third, teachers indicated that practical ideas would be instrumental in changing their attitudes toward and practices regarding creative dance because they could transfer these ideas into their specific teaching situations. Teachers also said that practical ideas related to integrating creative dance across the curriculum would be particularly helpful. Interventions should therefore feature explicit information on how to integrate creative dance into core programs, for instance, demonstrations of manual creative dance skills designed to enhance lessons in mathematics, science, language, and other content areas.

Fourth, teachers need access to more resources on creative dance. Intervention should therefore include providing teachers with resource lists of journal articles, textbooks, and teaching videos. Teachers expressed a strong wish for resources on how to present abstract ideas in concrete form. Addressing the issue of facilities by showing teachers how to implement creative dance with limited space is also necessary.

Finally, there is a definite need for departments of education to create support documents on creative dance. The effect of current priorities is to deemphasize the arts in general and creative dance in particular. For example, all 20 teachers reported that their ministry has produced no support document on creative dance, that there is little emphasis on the subject in the curriculum, and that creative dance has no priority on professional development days or in workshops. By contrast, documents do exist in other subjects, for example, science, drama, and music. There is a parallel need for departments and school boards to recognize professional development in creative dance and to provide relief time for such development (Guskey, 1986). Support and encouragement from all levels of administration seems basic to improving and changing the situation (Courtney & Park, 1981; Fullan, 1982; Goodlad, 1984).

Conclusion

This article explores why elementary school teachers are not currently implementing creative dance in their classrooms. The findings point to the need to provide support and inservice for elementary school teachers to address their needs in relation to attitudes and beliefs, knowledge or skills, and resources or facilities.

Understanding why creative dance is not more prevalent in the elementary schools is the first step in planning an appropriate intervention strategy. The findings discussed in this article also suggest that providing intervention addressing these expressed needs could go a long way toward ensuring that more children are exposed to this noncompetitive, nonthreatening activity.

References

Best, D. (1985). *Feeling and reason in the arts*. London: Allen and Unwin.

Boorman, J. (1973). *Dance and language experience with children*. Toronto, ON: Longman.

Courtney, R. (1980). *The dramatic curriculum*. London, ON: Faculty of Education, University of Western Ontario.

Courtney, R., & Park, P. (1981). *Learning in the arts*. Toronto, ON: Ministry of Education.

D'Houbler, M. (1957). *Dance: A creative art experience*. Madison, WI: University of Wisconsin Press.

Fowler, C. (1989). The arts are essential to education. *Educational Leadership Journal, 47*(3), 60-64.

Fullan, M. (1982). *The meaning of educational change*. Toronto, ON: OISE Press.

Gardner, H. (1985). *Frames of mind: The theory of multiple intelligence*. New York: Basic Books.

Goodlad, J.I. (1984). *A place called school: Prospects for the future*. New York: McGraw-Hill.

Guskey, T.R. (1986). Staff development and the process of teacher change. *Educational Researcher, 15*(5), 5-12.

Laban, R. (1968). *Modern educational dance*. New York: Praeger.

MacDonald, C.J. (1989). *Changing elementary-school teachers' attitudes and practices regarding creative dance*. Unpublished doctoral dissertation, University of Toronto.

Miller, J. (1988). *The holistic curriculum*. Toronto, ON: OISE Press.

Regan, E., & Leithwood, K. (1974). *Effecting curriculum change: Experiences with the conceptual skills project*. Toronto, ON: OISE Press.

Riley, A. (1984). *The interrelationships and effects of creative dance on the physical self-esteem, body image, and problem solving of grade four children*. Unpublished doctoral dissertation, University of Toronto.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 2, June 1991, 167-178*

GLENDA H. CAMPBELL-EVANS
Edith Cowan University, Perth, Australia

Nature and Influence of Values in Principal Decision Making

The research addressed five questions concerning the place of values in administrative decision making. This was done in an effort to improve understanding of the nature and influence of the values in professional decision making. Data were collected in three phases from elementary and elementary-junior high principals in a small, urban school district in Alberta. The phases included open-ended interviews, simulated decision tasks, and structured interviews. Beck's (1984a) value framework guided primary analysis and the work of Hodgkinson (1978) contributed to interpretation. Principals made multiple reference to three of Beck's five value categories: social and political, basic human, and moral values. Of these three, social and political and basic human were dominant. A primary set of 15 individual values emerged from the categories.

Discussion of administrative decision making typically focuses on a description of a rational view of the decision making process. This rational view separates value from fact and emphasizes the importance of the latter in decision making. Although such a view has generated models and systems of decision making, it does little to advance our understanding of the influences and rationale of the individual decision maker. That is, it provides some insight into the *what* of decision making but ignores the *why*.

The face of educational administration literature, traditionally sparse in its treatment of the nature and influence of values in administration, is changing. The work of Evers, 1985; Lakomski, 1987; Hodgkinson, 1978; and Greenfield, 1986 is evidence of this change. These authors have contributed to a debate and a growing body of literature that addresses the significance of values in theoretical terms. As suggested by Begley and Leithwood (1990, p. 338), "a theoretical perspective that accommodates the existence of values as influences on administrative practice at least potentially expands our understanding of administrative actions beyond that which is possible employing the exclusively rational frameworks normally associated with effective schools." This important first step provides the foundation for the empirical study of values in the administration of schools. Study that is focused on the decision maker contributes to an understanding of principal practices.

This research contributes to a group of studies attending to the causes of principal practices (e.g., Begley, 1988; Begley & Leithwood, 1990; Leithwood and Stager, 1989; Stager & Leithwood, 1988). In these works, internal influences or values are deemed to be an important cause of action. As

Glenda Campbell-Evans is a lecturer in the Policy and Administrative Studies specialization in the Faculty of Education. Her research efforts are currently focused in the area of the professional development of practicing and aspiring school administrators.

Holmes (1986, p. 89) suggests, the facts we perceive depend on the values we have and this interrelatedness of fact and value accounts in part for action. It is through values that individuals formulate a standard of acceptable behavior. This standard reflects a fit between facts and values, that is, one judges the facts from a values position. Values are expressed through actions and through one's judgment of the actions of others.

This study addressed questions concerning the place of values in administrative decision making in an effort to improve understanding of their nature and influence. The research focus was to highlight the value component of principals' decision making processes by identifying and categorizing those types of values employed so that values and their influence on the process could be examined in specific terms.

Specifically the research sought to answer five research questions. First, to explore the nature of values:

1. What types of values do principals use in making professional decisions?
2. What importance do principals attach to different types of values?
3. To what extent are values influential in a principal's decision making process as compared with other relevant influences?

Then, to investigate how values influence decision making:

4. Do principals with different orientations to the role bring different sets of values to bear on decision making?
5. How do principals respond to situations in which proposed courses of action conflict with their values?

The questions were addressed through three phases of data collection. The intent was to collect descriptive data and subsequently to use these data to try to understand the values component of decision making actions, choices, and priorities.

Framework

Literature on the topic of values indicates the existence of discrete categories or types (e.g., Beck, 1984a; Hodgkinson, 1978; Rokeach, 1973). A categorical organization of values allows for the investigation of preference for some value types over others. This preference may explain the variation between individuals' actions. That is, in the same circumstances different principals make different decisions and consequently act in different ways. The differences may be accounted for by the relative importance each attaches to particular values and value types within the value domain.

Facts are the *is* component of experience and they are interpreted through an individual's system of values, the *should* component. Facts provide an information pool that is important when an individual is faced with making a choice between alternatives: the essence of decision making. Inclusion of values in discussion of administrative decision making provides a more comprehensive description and understanding of the process, as well as recognizing their role in principal practice. Although this description attends specifically to the fact-value components, these are only two of a larger set. Berleant (1973), for example, believes that values must be conceptualized with "reference to the human activities of believing, striving, responding in all sorts of ways to the things and events men encounter" (p. 28).

Beck's (1984a) framework was used to conceptualize the nature of values. He suggests that a fairly common set of universal values exists. Within this common set, five general types can be identified. These types and examples of each are presented in Table 1.

Beck (1990) suggests that it is a mistake to attempt to deal with value issues without reference to the "large questions" of life. "We cannot resolve or even come close to grips with value issues without bringing to bear perspectives of this kind" (p. 11). Analogous to Beck's "large questions" is the organizational context of school. Singular decisions and actions must be made in relation to the big picture of school life and goals. Priorities and emphases may change and shift over time and with respect to specific circumstance. An integrated look at values, decision making, and principals that focuses on the interplay of these components within the framework of the organizing structure is important. The value component of principals' decision making must be explored within the context of school.

Recognition of the value component of organizational life is central to the work of a number of theorists. It was within this perspective that the research was framed. Greenfield (1985, p. 3), for example, sees values as the means to form and order facts. It is values, he contends, that structure decision making.

Decision making is primarily focused on choosing one course of action from a possible many. The choice is made within a context and is influenced by elements of the environment. Decision makers, therefore, are bounded by environmental restraints such as available time and resources as well as by the personal constraints of skill, knowledge, and processing capacity. As the administrator of a school, the principal engages in significant and varied instances of decision making. Leithwood and Stager (1986, p. 2) see the job of

Table 1
Examples of Value Types

Basic Human Values:
survival, happiness, companionship, friendship, helping others (to an extent), participation in community, self-respect, respect from others, knowledge, freedom, a sense of meaning in life, etc.
Moral Values:
carefulness, responsibility, courage, self-control, reliability, truthfulness, honesty, politeness, fairness, unselfishness, etc.
Social and Political Values:
peace, justice, due process, tolerance, participation, cooperation, sharing, loyalty, solidarity, citizenship, respectability, etc.
Intermediate-range Values:
food, shelter, entertainment, relaxation, fitness, good family relationships, ability to read, write and calculate, meaningful employment, financial security, etc.
Specific Values:
a car, a telephone, a particular neighbourhood, a particular friendship, a high school diploma, a good TV, a particular sport, a particular political party, a particular job, etc.

Source: (Beck, 1984a, p. 2)

a school principal as one where the individual is bound to make choices that best suit the context in which they find themselves.

It was within a framework where, first, both fact and value are seen to affect the importance and relevance assigned to various factors within each decision alternative and, second, where values provide structure and organization for the determination of action that this study was conducted. From this conceptual position it is possible to identify and explore the value component of decision making.

Method

This research was conducted with the principals of elementary and elementary-junior high schools in a small, urban school district in Alberta. There were eight such schools. One female and seven male principals were interviewed. Data were collected from each principal on three separate occasions in 1987. Total interview time per principal averaged six hours. All collection was undertaken by one researcher and was preceded by pilot interviews. The three phases of data collection included an open-ended interview (Phase 1), simulated decision tasks (Phase 2), and a structured interview (Phase 3).

Phase 1 centered around a question set for principals prior to the interview, that being: "Which of the professional decisions you have made during recent months would you identify as important?" In their discussion of the decision situation, principals were asked to address decision influences, actions, choices, and restrictions. Phase 2 decision tasks provided principals with a common set of decision situations. These simulated existing or previous District circumstances. It was intended that relevant tasks would allow principals to speak from experience rather than from speculation. This satisfies the requirements of content validity. The notion of relevance corresponds to the importance of content validity. "A simulation will have content validity if it includes the major relevant aspects associated with the dynamic decision making situation" (Thompson & Keon, 1982, p. 440). The purpose of the Phase 3 interview was to determine the research subject's orientation to their professional role. This phase utilized the profile work of Leithwood and Montgomery (1986). Responses to their structured Profile interview questions form a description of the individual's practice. The description is focused in the areas of goals, factors, strategies, and decision making.

All phases of data collection emphasized discussion. If our values and interests shape how we study and discuss reality, then discussion of that reality must reflect our values and interests. Exploring the realm of values is suited to a discussion approach. Through discussion, it was my intent to pursue the meaning behind action and come to an understanding of the principal's decision making in terms of the nature and influence of values inherent in the process.

All sessions were tape-recorded and transcribed. General themes and priorities evident in Phase 1 were shared with principals at the conclusion of Phase 2 for confirmation and verification. Principals validated the data during this playback exercise.

Analysis

Kluckhohn (1951, p. 395) defines value as "a conception, explicit or implicit, distinctive of an individual or characteristic of a group, of the desirable that influences the selection from available modes, means, and ends of action." He links values to action by identifying them as one determinant of that action. To identify the value component of reported action, interview statements were analyzed according to Beck's (1984a) value framework. Words such as should, ought to, like to, good, important, have to were clues to value statements. Use of such words indicated a sense of what was desirable for that individual and important to him or her. Analysis of the data, organized within the value framework, resulted in drawing conclusions concerning its content. By equating importance with frequency, a noninferential strategy was employed, that is, no judgment was necessary to establish importance.

To determine the dominant value types and the importance principals attached to each type and values within each type, the data were examined in a number of ways. These included: (a) the number of references principals made to a particular value was noted; (b) the number of principals making reference to particular values was determined; (c) dominant values in Phase 1 were compared with dominant values in Phase 2 to determine consistency between data collected in an open-ended decision making discussion (Phase 1) and an incident specific decision task (Phase 2); (d) any value identified as important by all principals was noted. From these four different analyses, patterns in the type and importance of values used by principals in professional decision making, emerged. Table 2 displays the results.

Data were also examined for the consistency of each principal's responses across phases and for values dominant in a particular decision task.

Results

Types and Importance of Values

The principals made multiple references to three of the five specified value categories in their school based decision making: basic human values, moral values, social and political values. Of these three, two were dominant. Within these value categories a set of 15 individual values dominated: seven social and political, five basic human, and three moral. These 15 form the primary set of data as represented in Table 2.

Frequency data relevant to Phase 1 and to Phase 2, along with an overall tally and corresponding ranking are recorded in column two. Following that, frequency and rankings are provided (column three) based on the playback data. The playback exchange served as both a perception check and a summary. It provided a means of verifying with each principal, my overall, general impression of him or her as a decision maker based on the Phase 1 interview. The Principal/Phase data displayed in column four, indicates the number of principals who reflect the values in the primary set in Phase 1 and again in Phase 2. Column five provides data concerning the pervasiveness of each value. This is done through noting the number of decision tasks in which the value was mentioned.

The meaning and use of each of the 15 values in the primary set is specified through comments made by principals during the collection of data.

Table 2
Set of Primary Values

I Values	II Frequency				III Play-back		IV Princ/Phase $n=8$		V # of Dec $n=5$
	Ph.I	Ph.II	+	Rank	Freq	Rank	Ph.I	Ph.II	Total
Basic Human									
knowledge	32	40	72	3	8	3	8	8	5
helping others	28	17	45	7	2	8	8	7	5
happiness	23	10	33	9	6	5	7	5	4
respect for others	45	16	61	4	9	2	8	7	3
survival	9	13	22	13	0	13	4	5	4
Moral									
responsibility	16	30	46	6	2	8	7	8	5
carefulness	8	17	25	11	2	8	5	6	5
fairness	9	11	20	14	1	12	5	6	4
Social and Political									
participation	61	53	114	1	10	1	8	8	5
cooperation	30	17	47	5	2	8	8	6	5
sharing	32	54	86	2	7	4	8	8	5
solidarity	17	6	23	12	5	6	7	5	3
respect for authority	13	32	45	7	0	13	7	8	5
commitment	18	9	27	10	4	7	7	7	4
community	12	8	20	14	1	12	3	6	3

Legend: II Frequency: # of mentions by 8 principals/phase
III Playback: # of reflections of a particular value in the playback description
IV Princ/Phase: # of principals who reflected particular value/phase
V # of DEC.Tasks: # of decision tasks in which particular value mentioned.

To clarify the meaning attached to each of these values, an illustration of each, within its category, is presented in column two of Table 3.

Values in the social and political and basic human value categories dominated the data. In terms of frequency, principals' comments most often reflected values of the social and political type (362), followed by basic human values (233), moral (91), specific (7), and intermediate-range (4). Within the possible range of values identified by Beck (1984a) it is notable that 15 values prevailed.

Principals referred to participation more often than any other value in Phase 1 and it was second to sharing in Phase 2, by only one tally. This value figured prominently in the playback and received mention in each decision task as well as by all principals during each phase. Overall, its dominance is clear. Principals indicated that they wanted to be involved with the activities of the school and that they attached importance to involving others. Sharing also had a particularly high profile in the data. It was the most frequent Phase 2 value mentioned by all principals and in each decision task. Overall, sharing holds the second ranked position with knowledge. Knowledge was

Table 3
Value Illustrations

I Values	II Illustration
Basic Human	
knowledge	- I think the knowledge is very, very important - another one I'm committed to is what Mannat calls density of curriculum
helping others	- it would be picking up for someone else who had problems and stresses - I would be prepared to work with him
happiness	- open school where everybody had a lot of fun so I have a vision that this is a nice place to come to and students want to be here
respect for others	- kids are neat people - value basic interest in people
survival	- we agreed on what we could bite off and chew reasonably - trying to do everything and be everything to all—just couldn't do it
Moral	
responsibility	- I think it's my responsibility as an administrator to look at these details ... those are administrative details that I think are my responsibility rather than the teachers - I think it's up to the principal, not on a direct basis, but very, very quickly to just move around and just make sure
carefulness	- we'd have to make sure we'd done our homework and that we had a very sound case - don't kid yourself, I'll be monitoring it. I'll let them try it but I will monitor it carefully
fairness	- I think that's where the fairness and the objectivity is—always in the foreground - its hardly fair to keep going and taking and never giving
Social and Policital	
participation	- we felt we needed to get our students, especially at the grade six level, more involved in the school - I still try and involve people who were concerned about the issue
cooperation	- hoping that we'd get the staff to trust each other and to work together to be able to work well with classroom teachers - if you're going to be successful and reach these goals, you have to work together to get there
sharing	- share ... with the community in handbook form - you bring it back, discuss it and share it
solidarity	- a school that really has an identity ... a school that has a sense of direction with a plan of action - I just wanted to focus on something the students could think about, that the teachers could think about, and that I could think about
respect for authority	- they (the Board) are the ones who make the rules, they set the standards for the district, we just implement it - if the policy is laid down early and the teachers are all aware of it and again it would be set down and then presented to the staff
commitment	- it's really important that they have ownership into it - it does take a group commitment
community	- help the overall staffing situation in the district and provide someone else with an opportunity - you have to look at the total picture and how it affects the greatest number of people

the dominant basic human value in Phase 2 mentioned by all principals in each phase and in each decision task. Knowledge is the highest ranked basic human value and third overall when the combined Phase 1 and 2 frequency is considered.

Influences of Values as Compared With Other Relevant Influences

Principals were asked to identify in a general sense the influences on their decision making. From their responses, a set of six primary influences emerged, three internal and three external. The internal influences of concern for students, effect of decision, and commitment of others to the decision are illustrated respectively in the following interview excerpts: "kids come first and then other things fall into place after that. So that's the strongest influence," and "in cases where not every staff member is influenced or affected equally, then more isolated discussion takes place with individuals who are affected rather than with those that are less affected," and "I think buying into an idea by the people being affected is."

Principals identified the external influences of time, money, and factual information. The importance of time is reflected in comments such as, "you could look at costs as well—what's its costing us ... It's costing us in terms of time." Money was identified as an important influence on the decision making process as evidenced by; "they have to know that dollars count." Comments such as, "you state the facts" and "we'd have to make sure we'd done our homework" and "I'd have to gather a lot of information," indicate the attention given to facts and reflect the presence of the influence.

Situations of Conflict

A focus on conflict situations provided another opportunity to explore the influences values play in decision making. When principals find themselves in a conflict situation, how do they resolve it? The conflict arises as a result of external or internal directives or pressures. External conflicts are those that arise because an external directive contradicts a value held by the principal. The data indicated that when policy conflicted with their values, principals responded initially by implementing that policy. They dealt with or resolved the conflict by moving to change or amend it to be more compatible with their personal value position. Such external pressures are assessed and courses of action determined through the principal's individual value filter.

In the case of internal conflicts, that is, when principals are uncertain of how to proceed, principals are not clear as to which values ought to be guiding their action. The eight principals recalled situations representative of internal conflicts by discussing situations when "their actions had gone against their better judgment." In these cases the value choices needed to precede action.

Principals' values give meaning to information that is encountered. Values, therefore, are influential in a principal's decision making process in comparison with other influences. When, through the values filter, information is interpreted as being important, that information becomes an influence on principals' thinking and possibly on their action.

Orientation to the Role

The traits specified by principals in explaining the differences between the administration of their school and others in the district related to the principal's orientation as determined by the *Principal Profile* (Leithwood & Montgomery, 1986). The Profile describes principal practice in terms of four orientations; Administrator, Humanitarian, Program Manager, and Problem Solver. A value-behavior relationship existed. In explaining the differences in the administration of different schools, principals highlighted those aspects of school life specific, important, or unique to their school. These specifics were a result of principal decision making and action. Principals indicated that they consciously noted these differences and strove to highlight the factor that identified their school. Their actions promoted the unique features.

The distinguishing feature of the school was consistent with the principal's orientation in seven of the eight cases. Knowledge, responsibility, and participation were dominant values for the Program Manager who, as the name implies, has a program focus. The consistency in these data reflects a singular and focused image of principal with this orientation. The dominant basic human value of knowledge seems logically consistent with the priorities of the Program Manager perspective. Similarly, in the case of the second represented orientation, that is, Humanitarian, the basic human value of respect for others was dominant. Although too limited to be definitive, the data indicate a relationship between a principal's orientation and his or her dominant values that consequently influence action.

Discussion

This study examined the nature and influence of values in principal decision making using Beck's (1984a) values framework to guide analysis. The eight principals appeared to attach particular importance to a set of 15 values that span three value types: basic human, moral, and social and political.

Why is it that these elementary school principals discuss school decision making in these terms? Basic human values are, as the name implies, basic, that is, linked to human need. These values are seen as important in themselves and also in relation to values in other categories. Principals' comments indicated a similar or common need to treat people (staff, students, parents) in a manner that communicated the principals' view that they had importance. It is reasonable, therefore, that reference to this type of value would punctuate principals' speech. Basic human values are part and parcel of everyday living.

Similarly, social and political values were frequently identified as being particularly important to the principals throughout the data analysis. Social and political values are social in nature, and these values of consensus and consequence have relevance for group settings. The collectivity or community is a necessary condition for the reflection of these values and the nature of school satisfies that condition. School is a social organization. Learning and practicing socialization skills is important in the school environment. The community setting necessitates attention toward participation, cooperation, sharing, solidarity, and so on. The values of cooperation,

sharing and participation received repeated principal attention. Perhaps this is so because these are not mutually exclusive. Participation is necessary in order to cooperate and sharing may aid cooperation. In order to share time, ideas or materials, one is cooperating with others and in order to participate, one may need to cooperate. Consequently, the socially oriented values logically play an important role in school life and were dominant in the data.

Hodgkinson (1978) discusses consequences and consensus values in social, rational, and reasoning terms. For Hodgkinson, they are "values associated with custom and reason" (p. 145) and "the grounds are social for they depend upon collectivities and collective justification" (p. 112). Schools provide instances of the social context to which Hodgkinson refers. The importance of consequences and consensus is evident in the principals' attention to the effect of their decisions (consequences) and their preference to have the participation and commitment of staff (consensus). The decision making themes that emerge from the data exemplify principals' concern for consequences. Principals consider the consequences of decision making prior to the action.

Results of this study correspond to findings presented in similar research. In a study of computer adoption and use, Begley (1988) explored the influence of principals' beliefs and values on implementation. Values grounded in consequence and consensus were dominant. Using Hodgkinson's (1978) value model to guide analysis, Begley reported that, "The perceptions of most principals about computers seemed to be grounded predominantly at the level of consequences and consensus ... Few principals among the 15 interviewed were motivated by values of preference ... Values of principle were not frequently expressed" (p. 64). Application of these two perspectives resulted in analysis more similar than different.

Conflict, as an influence on decision making, was explored in both Phases 1 and 2. These data support the value position outlined in the conceptual framework of this study. The value system is fluid and dynamic, not static and set. When a principal described his or her action in a conflict situation in a way that reflected respect for authority, all other values important to that individual were not extinct; they assumed secondary positions. If values were an all-or-nothing phenomenon, then respect for authority, being stronger than all others in this case, would dictate behavior. The situation would stop at the initial "policy-following action" of the administrator. Principals described, however, that once the policy had been adhered to, they began to deal with the conflict. They all proceeded to "do something." Here, then, value structure shifted and adjusted and other values came to the fore. Sharing the situation, inviting participation, and seeking information followed implementation of the policy. These principal actions revealed attention to, and importance of, values other than respect for authority. That initial value, still in existence, then took a less prominent and influential role in the subsequent action of principals.

Research Directions

Specific attention to principal goals and the value component of those goals specified in Beck's and Hodgkinson's terms, is an important focus for future

research. The relationship of goals to values is one area for exploration. Beck (1984b, p. 8) suggests that although values are both means and ends oriented, basic human values tend to be more ends focused. Basic human values were prominent in the data; they had importance for principals. If values in this category are more absolute than instrumental, then they must represent the goals, not the processes of principals. If goals are value linked, perhaps ends focused basic human values can be explored through principals' public and private goals.

The Principal Profile (Leithwood & Montgomery, 1986) suggests that clearly stated principal goals contribute to improved student outcomes. Helping principals learn to clearly state goals and then looking at the influences of a clearly stated value stance and school vision on teacher and student behavior may aid our understanding of the importance of value awareness and its relationship to student outcomes. The data suggest a relationship between principal orientation and basic human values.

An examination of administrators' own process of decision making and attention to the roles values play in that process can occur through a principal's review of his or her recent school decisions. Recognition of similarities, themes, and consistencies would put principals in tune with their own decision making preferences of which they may or may not be aware. Attention to differences and inconsistencies may expose important influences, pressures, and demands in those particular decision making situations.

Knowledge of the role of values in principal behavior is still sketchy. There is no clear evidence or specification of the role, but existing research and theoretical writing have increased acknowledgment and awareness of the importance of value questions. This awareness can develop in a reflective direction. Reflection has application when values are recognized as an organizer in confused or complex situations and problems. In daily practice Schön's (1983) concept of reflection-in-action makes sense. In practice then, administrators need to start thinking about their values as they act and also to think about the role and influences of values after they have acted. In order to do this effectively, professional development sessions must provide administrators with appropriate background and basis for the task.

References

Beck, C.M. (1984a). *The nature of values and implications for values education*. Unpublished Paper.

Beck, C.M. (1984b, November). *Our faith confronts differing life styles and value systems*. Presented to the Islington United Church School of Religion.

Beck, C.M. (1990, March). *"Faith" in a broad sense as a goal in promoting human development*. Paper presented at the Philosophy of Education Society Conference, Miami.

Begley, P.T. (1988). *The influence of personal beliefs and values on principals' adoption and use of computers in schools*. Unpublished doctoral dissertation, University of Toronto.

Begley, P.T., & Leithwood, K.A. (1990). The influence of values on school administrator practices. *Journal of Personnel Evaluation in Education, 3*, 337-352.

Berleant, A. (1973). Experience of values. In E. Laszlo & J. Wilbur (Eds.), *Value theory in philosophy and social science* (pp. 24-37). New York: Gordon & Breach Science Publishing.

Evers, C.W. (1985). Hodgkinson on ethics and the philosophy of administration. *Educational Administration Quarterly, 21*, 27-50.

Greenfield, T.B. (1985, May 28). *Putting meaning back into theory: The search for lost values and the disappeared individual*. Paper presented at the annual meeting of the Canadian Society for the Study of Education, Montreal.

Greenfield, T.B. (1986). The decline and fall of science in educational administration. *Interchange, 17*(2), 57-80.

Hodgkinson, C. (1978). *Towards a philosophy of administration*. Oxford: Blackwell.

Holmes, M. (1986). Comment on Greenfield. *Interchange, 17*(2), 80-90.

Kluckhohn, C. (1951). Values and value-orientations in the theory of action: An exploration in definition and classification. In T. Parsons & E.A. Shils (Eds.), *Toward a general theory of action* (pp. 388-433). Cambridge, MA: Harvard University Press.

Lakomski, G. (1987). Values and decision making in educational administration. *Educational Administration Quarterly, 23*(3), 70-82.

Leithwood, K.A., & Montgomery, D.J. (1986). *The principal profile*. Toronto, ON: OISE Press.

Leithwood, K.A., & Stager, M. (1986, April). *Differences in problem-solving processes used by moderately and highly effective principals*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Leithwood, K.A., & Stager, M. (1989). Expertise in principals' problem solving processes. *Educational Administration Quarterly, 25*(2), 126-161.

Rokeach, M. (1973). *The nature of human values*. New York: Free Press.

Schön, D.A. (1983). *The reflective practitioner*. New York: Basic Books.

Stager, M., & Leithwood, K.A. (1988). *Flexibility in principals' problem solving processes* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Thompson, K.R., & Keon, T.L. (1982). The creation of certain and uncertain environments: A study using a management simulation. *Simulation & Games, 13*(2), 437-450.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 2, June 1991, 179-192*

W. TODD ROGERS
University of Alberta

Educational Assessment in Canada: Evolution or Extinction?

Educational assessment in Canada is not, nor will it become, extinct. Accountability testing, with its movement toward new methodologies, is here to stay. And it is equally clear that student assessment is an integral part of teaching and learning, and that good teaching cannot exist without valid assessment of the students to be taught. The evidence suggests, however, that teachers do not perceive themselves, nor do they appear to be, competent in classroom assessment. Further, it appears that teacher education programs do not provide adequate training in classroom assessment. To address this issue, five suggestions, derived from course outlines obtained from 33 universities across Canada and from the literature, are made together with a call for a review of the assessment component of the preservice education program at each Canadian university in terms of the Standards for Teacher Competencies in Educational Assessment of Students.

> The elemental stuff of the measurement courses and textbooks is as foreign to classroom teachers as spectral analysis of glaze mixtures is to potters—and apparently as irrelevant. (McLean, 1985, p. 39).

There appears to be a paradox in educational measurement today. While assessments of achievement are being more urgently called for and more widely employed than ever before, those responsible for training teachers are seemingly turning away from including measurement and evaluation as a major component of their preservice teacher education programs.

Three reasons for this apparent paradox are suggested in this article. First, the dominant view of measurement over the last 20 or 30 years on accountability testing has unintentionally contributed to widening the gap between measurement and classroom instruction. Second, while teachers spend a great deal of their time on classroom assessment, evidence suggests that teachers do not perceive themselves, nor do they appear to be, competent in classroom assessment procedures. This lack is ascribed to the observation that teacher education programs do not provide adequate training in classroom assessment. The article then concludes with a series of recommendations that, hopefully, will serve as a beginning to improving classroom assessment practice.

The Dominant View of Testing: Accountability Testing

The trend toward new and expanded testing reflects, in part, a call for greater accountability. Much of the public concern through the 1970s and into the 1980s focused on "declining standards" or on related matters indica-

Todd Rogers is a professor in the Department of Educational Psychology and Director of the Centre for Research on Applied Measurement and Evaluation. He specializes in measurement and evaluation.

tive of the fear that schools were not doing a good job of quality control. Politicians responsible for allocating funds for education wanted to know if they were getting their money's worth.

Consequently, large education reforms implemented during the past few years have explicitly incorporated either the use of new tests or the expanded use of existing ones. In Britain, for example, new national tests are a key tool in the government's effort to enact far-reaching changes in that nation's educational system. Central to this reform is a unified national curriculum backed by a series of tests administered to children 7, 11, 14, and 16 years of age. In the United States, the governing board of the National Assessment of Educational Progress recently announced a set of uniform standards to be applied to NAEP reports. Intended to further state-by-state comparisons and to provide an incentive to improve achievement, the standards are to remain constant for the foreseeable future so that results of each of its biannual assessments can be compared (Barringer, 1990; Linn, 1989).

Turning to the situation in Canada, concerned with the quality and effectiveness of their education systems, the Council of Ministers of Education, Canada will assess the achievement of 13- and 16-year olds across the country in literacy and numeracy in 1992. Several provinces (e.g., British Columbia, Alberta, Ontario) are looking at ways to incorporate performance measurement and portfolio assessment into their provincial and assessment programs. Interested in how they fare internationally, nine provinces are taking part in the Second International Assessment of Educational Progress in 1991. All provinces will participate in the Third International Study of Mathematics and Science scheduled for 1993/1994.

Available evidence suggests that over the past 20 or so years this call for accountability has influenced measurement practice in a major way: The dominant view regards measurement as the means of documenting student achievement by using externally developed standardized paper-and-pencil tests, including provincial assessment instruments and examinations, for public accountability (Stiggins, Conklin, & Bridgeford, 1986). This view may have unintentionally contributed to widening the gap between measurement and instruction. Glass (1986) laments,

> In short, it [testing] sold out to the highest bidder; it went Establishment. In the world of work and in education, testing stepped forward to play the role of gatekeeper and management tool in the processing of human lives for the meritocracy: professional lighting rod attracting and grounding the anger of the excluded and discarded (p. 14).

Distinction Between Accountability Testing and Classroom Assessment

In advancing the premise implied in the previous paragraph that measurement for instructional and grading purposes ought to occupy the preeminent position, it is not intended, nor should it be construed to mean, that testing programs initiated and conducted in Canada by ministries of education and local school boards should be ignored. Indeed, as pointed out by Traub (1990), governments that have public support for "external surveillance of student achievement" will continue with provincially [and district] mandated testing programs (p. 91). But what needs to be understood and preserved are the

distinctions between accountability testing and classroom assessment (Cole, 1988; McLean, 1985, 1990; Shepard, 1989). Briefly, as displayed in Table 1, measurement to serve accountability and policy goals typically is mandated by an external (to the classroom) agency marked by a high pressure environment, typically highly reliable, widely applicable, sensitive to aggregated change, and centrally processed. Results are of a summative nature and must be in a form useful to policy makers, which usually means reducing complexity to, preferably, a single score for a school, district, or province. In contrast, assessments designed to support instruction and to be used in grading are teacher mandated (although within reporting constraints typically set outside the classroom (see Wilson & Rees, 1990), take place in a somewhat lower pressure environment, incorporate both formal and informal techniques and strategies, less reliable (although aggregated scores are more reliable), adapted to the local (classroom) context, sensitive to short-term change in students' knowledge and understanding, and locally scored. They provide immediate, detailed, and complex feedback at the student level, and they incorporate tasks of instructional value themselves.

The Nature and State of Measurement and Classroom Assessment

Research on teaching reveals that day-to-day assessment of student learning is unquestionably one of the teacher's most demanding, complex, and important tasks. Every model of the teaching/learning process requires that teachers base their decisions—instructional, grading, and reporting—on some

Table 1
Distinction between Accountability Testing and Classroom Assessment

Accountability Testing	*Classroom Assessment*
Serve accountability and policy goals	Serve instruction and grading
Provincially, district mandated	Teacher mandated
High pressure environment	Lower pressure environment
Formal	Formal and informal
Reliable	Less reliable[a]
Widely applicable	Adapted to local context
Sensitive to "accumulated" change	Sensitive to short-term change
Centrally processed	Locally scored
Aggregated (school, district, province) reporting	Student reporting
Summative	Formative and summative
Reduce complexity to a single score	Detailed and complex feedback
	Incorporate instructional tasks

Adapted from Cole (1988), McLean (1985, 1990), and Shepard (1989)
[a]Although individual teacher made tests and assessment instruments are likely less reliable than a standardized test, the accumulation of data gathered about individual students during the course of a reporting period has greater accuracy.

knowledge of student progress toward desired learning outcomes. The complexity of classroom assessment—the process by which inferences are drawn about student progress—is seen from the range and frequency of teachers' decisions and the plethora of student learning outcomes and behaviors teachers consider when making these decisions.

Yet, given the prevalence and importance of classroom assessment, our understanding of in-class assessment practice still needs to be more fully developed. Following their review of research on testing in schools, Lazur-Morris, Polin, May, and Barry commented in 1980 that,

> In-class assessments made by individual teachers have yet to be examined in depth. How these and other assessments are united with teacher instructional decision-making processes and how they affect classroom organization and time allocation to other objectives are areas that should be explored. Teachers place greater reliance on, and have more confidence in, the results of their own judgments of students' performance, but little is known about [these] kinds of activities. (pp. 24-25)

Since that time, several research studies in which teachers and students were surveyed and interviewed and/or classrooms were observed have revealed a series of consistent findings. The findings of these studies are summarized in Table 2. Taken as a whole, they confirm the earlier suggestion that the day-to-day assessment of student learning is indeed a complex and demanding task. However, while teachers appear to value classroom assessment as an instructional tool and feel assessments benefit their students, the formative purposes give way to summative purposes with increasing grade. As described by Wilson (1990), the use of assessment information changes "from internal classroom uses to external public ones" (p. 16). Teachers work alone when developing their assessment methods. Although they use a variety of techniques, they tend to prefer objective, short answer/completion paper-and-pencil tests. Further, they tend to focus their items on lower order knowledge and skills; their tests provide little indication of the attainment of higher order cognitive knowledge and processes.

Teachers appear reluctant to commit the time needed to evaluate the quality of their own tests, and there is strong evidence to suggest they are ill-prepared to engage in quality control. The evidence further suggests that teachers' grading practices are suspect. Student grades are based more on the assessment of lower cognitive skills and less on desired higher level thinking. Frequently factors other than those of interest are included in a grade. And, while teachers suggest that by taking the simple sum of the components to be used to form the final score each component will count equally, differences in the standard deviations of the component scores suggest otherwise. As Messick (1989) points out in his excellent chapter on validity, problems such as these serve to undermine valid score interpretation and use.

To illustrate these issues, consider the following two examples. The first demonstrates the possible confounding of score interpretation due to the presence of construct-irrelevant test variance (Messick, 1989, p. 34). Test-wiseness is a cognitive ability or set of skills which a test taker can use to improve a test score no matter what the content area of a test (Benson, 1988;

Table 2
Summary of Findings from Research on Classroom Assessment

1. Teachers view tests, particularly teacher-made tests, as important instructional tools. They report tests
 - help focus teaching
 - provide feedback on instruction
 - provide feedback on student progress.
2. Teachers believe tests are of direct benefit to students
 - act of taking a test is a learning experience
 - motivate students to study
 - create student competition
 - improves student interaction.
3. Teachers use tests, and argue it is better to give more frequent short tests than to give long tests infrequently.
4. The typical teacher spends at least 20% of his or her time on testing activities.
5. Teachers view themselves as being in charge of the testing done for instructional purposes, although external regulations and expectations colour much of the data gathering activities. The actual role of measurement and evaluation varies across grade level.
 - primary level—teachers tend to use classroom assessments to comment on the growth of individual students and the success of a particular unit or activity.
 - beyond primary—increasing emphasis given to producing marks for grading.
 - senior level—focus on external, end-of-year examinations (where they exist) and agreement between teacher and external examination marks.
6. Teachers indicate preference for creating their own test items, with the higher the grade level, the greater the tendency to use their own assessment instruments rather than published tests. They do use other sources, particularly "textbook publisher prepared items" and, in the case of external year end examinations, "old" exams.
7. Teachers use a variety of assessment techniques (paper-and-pencil tests, anecdotal records, published tests, laboratory write-ups, projects, essays, individual work contracts, oral tests). Teacher-made objective tests, particularly short answer/completion, are the most common.

 Assessment procedures differ as a function of grade level: Elementary, particularly primary, teachers employ observation of classroom work whereas secondary teachers prefer paper-and-pencil tests. Formative purposes of evaluation at the elementary level give way to summative ones at the secondary level. Assessment procedures also differ as a function of subject area: mathematics and science teachers rely more heavily on paper-and-pencil tests. Teachers of speaking and writing tend to include more performance assessments (orals, essays, projects).
8. Teachers use essay tests to measure higher level cognitive skills. They believe essay tests
 - better evaluate higher order cognitive learning
 - in general provide a better evaluation of student learning.
9. Teachers tend to focus on lower order thinking skills (recall of declarative or procedural knowledge).
10. Teachers do not vary their assessment methods very much as the purpose of assessment varies. Nor to they anticipate having to change.
11. Teachers score and grade their own tests.
12. Teachers do little in the way of formal test analysis. They indicate they lack the time for effective quality control over the assessment instruments they develop. Further, they appear to lack the requisite skills.
13. Grading practices of teachers appear to be faulty. Academic grades incorporate other factors; quality of data used to form the grade not known; inclusion of formative and summative data; errors in aggregating scores.

Source: Anderson (1989); Anderson, Muir, Bateson, Blackmore, & Rogers (1990); Bateson (1990); Gullickson (1985, 1990); Gullickson & Ellwein (1985); Impara, Divine, Bruce, Liverman, & Gay (1990); Manke & Loyd (1990); Nickerson (1989); Stiggins & Bridgeford (1985); Stiggins, Conklin, & Bridgeford (1986); Stiggins, Frisbie, & Griswold (1989); Wilson (1990).

Millman, Bishop, & Ebel, 1965; Sarnacki, 1979). If a test taker possesses test-wiseness—partial knowledge about what an item is measuring and the ability to take advantage of item clues or cues—and if the examination contains flawed items with clues or cues, then the combination of these two factors can lead to an improved score. In contrast, students low in test-wiseness will be penalized every time they take a test that includes test-wise components.

Table 3 contains a summary of an analysis of the influence of test-wiseness on student performance on six grade 12 provincial examinations administered in British Columbia in June 1989 (Rogers & Bateson, in press). As shown, 43 to 80% of the items on these six examinations, which were developed by committees of experienced teachers, contained test-wise cues. Further, the results reveal the mean performance of the students on the subsets of items identified as test-wise susceptible exceeded the mean performance on the subset of nonsusceptible items by 9.2 to 16.4% ($p < .01$). These findings seriously suggest that interpretation of examination scores as valid measures of student knowledge and understanding of the subject matter and processes to be measured may be unwarranted. To the extent the partial knowledge is considered not relevant to the achievement tested, the interpretation of a test score will be confounded by construct-irrelevant easiness (Messick, 1989, p. 34). But if this partial knowledge is considered relevant, then the interpretation of the total score as a valid indicator of achievement may be more justified. Thus a potential validity problem exists when one attempts to interpret the meaning of the test score.

The second example illustrates the incorrect belief that by taking the simple sum of components (quizzes, weekly tests, etc.) to form a total score, each component will be counted equally in the total. Differences in the standard deviations of the component scores suggest otherwise. A sample of 31 elementary school teachers (Dauncey, 1986) was asked to regrade their students using total scores corrected in accordance with Ministry of Education guidelines (Taylor, 1979) for the unequal standard deviations of the component parts. In forming the initial total scores, the teachers believed they were counting the parts equally. The letter grades based on adjusted scores were then compared with the initial letter grades assigned by the teachers. A summary of the differences classified according to magnitude

Table 3
Summary of Test-wiseness Analysis

		Susceptible Test-wise Items			*Nonsusceptible Test-wise Items*		
Subject	*n*	*k*	$\bar{X}$	*SD*	*k*	$\bar{X}$	*SD*
English 12	736	21	67.9%	13.8%	7	58.8%	19.7%
Algebra 12	267	25	62.5%	15.6%	33	49.1%	14.7%
Geography 12	166	43	59.5%	11.8%	16	43.1%	14.0%
History 12	208	39	71.7%	16.7%	10	58.2%	20.1%
Biology 12	261	37	64.6%	15.6%	15	50.8%	18.1%
Chemistry 12	174	28	69.6%	15.4%	20	59.8%	17.6%

and direction of "change" is reported in Table 4. For example, a change from an initial letter grade of B to an adjusted score letter grade of C+ is recorded as a one-letter grade decrease. A change from a C+ to an adjusted letter grade of A is recorded as a two-letter grade increase. Almost half (46.2%) of the letter grades changed by at least one letter grade. How would a parent of one of the four children with a three-grade increase feel? Would you be satisfied with the initial grade? Would you feel comfortable, as a teacher, giving the parents of one of these children the initial grade?

Preservice and Inservice Teacher Education in Canada

Practicing schoolteachers suggest that the assessment of student learning is one of their most difficult tasks and that they have need of assistance. To illustrate, Marshall Arlin (personal communication, May 23, 1990) concluded his description of why an inservice university course was offered in response to a request from two groups of experienced primary school teachers with the following statements:

> These sections of the course were an attempt by the EPSE [Educational Psychology and Special Education, University of British Columbia] department to assist primary teachers in the implementation of the new Provincial Primary Program. Teachers had indicated that measurement issues were one of the hardest parts of implementing the new program.

Although sketchy and incomplete, the estimated number of inservice teacher workshops related to classroom assessment conducted during 1989/1990 in British Columbia (90; D. Bateson, M. Marshal, A. Taylor, personal communication, May 3, 1990) and Alberta (55; G. Thomas, personal communication, May 1, 1990; B. Krumbil, personal communication, May 8, 1990) portray a picture of need. Alberta Education received a similar indication of need in its survey of reactions to proposed changes to the provincial testing program at grades 3, 6, and 9. An overwhelming number of respondents (teachers, administrators, parents) believed it desirable that the Department provide inservice to assist teachers, school administrators, and trustees interpret and use achievement information (Horvath, 1990, pp. 2,4).

Table 4
Frequency of Letter Grade Changes

	f	%
One-letter grade increase	238	18.1
One-letter decrease	313	23.8
Two-letter grade increase	28	2.1
Two-letter grade decrease	23	1.8
Three-letter grade increase	4	0.3
Three-letter grade decrease	1	0.1
No change in letter grade	707	53.8
Total	1,314	100.0

Researchers who have assessed teachers' knowledge of classroom assessment practices have consistently voiced concern that teacher knowledge and skills are indeed inadequate (Bateson, 1990; Goslin, 1967; Gullickson, 1986, 1990; Mayo, 1964; Rudman et al., 1980; Stiggins, 1990; Stiggins, Conklin, & Bridgeford, 1986). Stiggins (1990) concludes: "The assessment community knows that teachers and administrators are in desperate need of assessment training; teachers and administrators know they need assessment training; but policy makers often are not aware of this gap in professional preparation" (p. 95).

One implication of these studies, and of those from which Table 2 was developed, is that preservice teacher education instruction in the area of classroom assessment is inadequate. There appear to be two plausible reasons for this: (a) many prospective teachers do not take a measurement course that deals with classroom assessment practices as part of their undergraduate teacher preparation; and (b) undergraduate measurement courses focus on the wrong topics or otherwise do not emphasize topics and activities related to classroom assessment.

Measurement and Evaluation Course Enrollment

To test the veracity of the first reason, representatives from the faculties/colleges/schools/departments of education (hereafter called faculties) at 33 universities across Canada were contacted by telephone and asked to describe the measurement and evaluation component of their teacher education program. The information and course outlines they provided revealed what may best be described as a patchwork of assessment training at the undergraduate level for both elementary and secondary teacher education students. Clearly, there is no uniform policy regarding the requirements for the assessment component of teacher education programs within provinces (with more than one university) or across Canada. Of the 33 faculties contacted, 25 offer at least one measurement course; eight offer none. Where more than one course is offered, it is either to serve different language groups, different levels of schooling, or different subject areas. Completion of the same measurement course is *required* of all elementary and secondary teacher education students in 10 faculties, of only elementary education students in five, of only postdegree elementary education students in one, and only vocational educational students in another. Ten faculties offer the same measurement course as an *elective* to both their elementary and secondary education students; an additional seven faculties offer measurement as an elective to only their secondary education students. The percentages of students who avail themselves of these electives vary. Across the 10 faculties with common courses, these percentages range between 5% and 34% for elementary, and 10% and 50% for secondary. The corresponding percentages for the secondary only option faculties for which information was available are 10%, 25%, 80%, and 90%.

These data are even more alarming when considered in light of the total number of teacher education students. It is not difficult to imagine that approximately 60% of the prospective elementary school teachers and up to three quarters of future secondary teachers will *not* complete an under-

graduate measurement and evaluation course prior to beginning to teach. Contrasted with the time these teachers will spend engaged in classroom assessment activities, this is a troubling image.

Students at both the elementary and secondary education levels do receive some measurement instruction in other undergraduate education courses that they are required to take. These courses may be found in curriculum, teaching methods, educational psychology, and/or foundations. The percentage of time devoted to assessment topics in these courses varies not only across faculties but also within faculties. Based on the information provided by the contributors, the time spent on measurement topics in these courses is equivalent to something between 10% and 50% (accumulated across courses where students take more than one) of a full-term (13 weeks) course. Further, disturbing as they were, contributors' comments left the distinct impression that the quality of measurement instruction is very much a function of the instructor's commitment to and his or her expertise in measurement.

Assessment Course Topics

Some investigators have argued that perhaps undergraduate educational measurement courses have focused on the wrong topics. In particular, elementary and secondary school teachers do not attach high importance to standardized tests and statistics, while professors teaching measurement do (Gullickson, 1986; Gullickson & Ellwein, 1985; Mayo, 1964). Further research (Gullickson, 1985; Salmon-Cox, 1982; Stiggins et al., 1986; Wilson, 1990) also suggests that teachers spend a great deal of time and energy in nontest evaluative activities. Gullickson (1985) found, for example, that of 50 topics rated for desired emphasis in a preservice measurement course, the greatest discrepancies between professors and teachers emerged on the following topics: Teachers wished for "great emphasis" to be placed on ways of "interviewing pupils and parents," "observing pupils' work habits," and evaluating "class discussions" and "interpersonal relationships"; the professors rated these topics as deserving only "slight emphasis." Indeed, the professors rated calculating the mean, variance, and correlation coefficient the highest.

Put in simple terms, teachers are more concerned with day-to-day issues related to the application of assessment processes and less with fundamental measurement principles. They want and need answers to questions like: What assessment procedure should I use for this purpose? How do I prepare for or construct this procedure? How do I score it? How do I use the information to make decisions?

In response to this situation and out of concern that potential educational benefits of student assessments be fully realized, a committee comprised of representatives from four educational professional associations in the United States (National Council on Measurement in Education, American Association of Colleges for Teacher Education, American Federation of Teachers, and National Educational Association) released in May 1990 *Standards for Teacher Competence in Educational Assessment of Students*. The Committee defined assessment

as the process of obtaining information which is used to make educational decisions about students, and to judge instructional effectiveness and curricular adequacy. The various assessment techniques include, but are not limited to, formal and informal observation, qualitative analysis of pupil performance and products, paper-and-pencil tests. Assessment competencies included here are the knowledge, skills and values critical to a teacher's role as educator. (National Council on Measurement in Education, 1990, p. 1).

Table 5
Appropriateness of Use of Standards for Teachers Competence in Educational Assessment of Students in Canada
1. University professors and Ministry officials ($n = 39$)
2. School teachers and administrators ($n = 124$)

			Reaction				
	Standard	*Group*	*SD*		*U*		*SA*
1.	Teachers should be skilled in choosing assessment methods appropriate for instructional decisions	1	0.0%	0.0%	2.5%	15.4%	82.1%
		2	0.0	.8	1.6	21.0	76.6
2.	Teachers should be skilled in developing assessment methods appropriate for instructional decisions	1	0.0	10.3	10.3	28.2	65.6
		2	0.0	.8	6.4	34.7	58.1
3.	The teacher should be skilled in administering scoring and interpreting the results of both externally produced and teacher produced assessment methods.	1	2.5	7.7	15.4	25.6	48.7
		2	0.0	2.4	5.6	40.3	51.6
4.	Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement	1	0.0	2.6	2.6	20.5	74.4
		2	0.0	.8	1.6	19.4	78.2
5.	Teachers should be skilled in developing valid pupil grading procedures which use pupil assessments.	1	0.0	2.6	10.2	20.5	66.7
		2	.8	2.4	14.5	28.2	54.0
6.	Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.	1	0.0	0.0	5.1	7.7	87.2
		2	0.0	0.0	2.4	17.7	79.8
7.	Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information	1	0.0	2.6	0.0	7.7	87.2
		2	0.0	0.0	8.9	28.5	62.6

Are these *Standards* appropriate for teachers in Canada? Shown in Table 5 is a summary of the degree to which a group of university education professors ($n=34$) and ministry of education officials ($n=5$) and a group of elementary and secondary school teachers ($n=95$) and administrators (n=29) agreed that the set of standards was appropriate to guide the preservice and inservice preparation of teachers. The university professors were surveyed at the beginning of the session dealing with the topics discussed in this paper and which was held during the 1990 conference of the Canadian Society for the Study of Education.[1] The teachers and administrators were from 29 public schools located in a large western Canadian city. They were surveyed at the beginning of the opening plenary session of a two-day conference on classroom assessment held in October 1990. As revealed in Table 5, a clear majority of respondents in both groups agreed that the Standards are appropriate to guide the assessment component of teacher preparation programs. More than 90% of the respondents in both groups agreed Standards 1, 2, 4, 6, and 7 were appropriate; in excess of 80% endorsed standard 5. Approximately 92% of the teacher/administrator group agreed standard 3 was appropriate; 75% of the university/ministry of education group did likewise. University professors who indicated that they were undecided about the third standard suggested that while teachers need to be able to administer, score, and interpret the results from teacher-produced assessment methods and externally produced tests such as provincial assessments and high school leaving examinations, they need not be skilled in administering, scoring, and interpreting externally produced, individually administered tests used for screening and diagnostic purposes.

Clearly, this acceptance of the *Standards for Teacher Competence in Educational Assessment of Students* presents a serious challenge to teacher educators as they design and approve programs for teacher preparation. The suggestions presented below, which were derived from the course outlines in the telephone survey and from the literature, are intended to help facilitate the changes in the assessment component of undergraduate teacher education programs so that prospective teachers will be adequately prepared for the classroom assessments they, as practicing teachers, inevitably will do.

1. Generally, interpretation and grading are two of the last topics listed in the outlines. Given the frequent problem of time running out, it is more likely than not that these two topics, rather than others, are given short shrift. Teachers, though, remind us of their importance. Steps must be taken to ensure both are adequately covered. For example, attention must be paid to the melding *evaluative* evidence so strongly desired by parents with *explanatory* evidence necessary to support a meaningful interpretation of that evaluative evidence.
2. Nontest procedures are typically stressed more in courses or sections of courses intended for elementary, particularly primary, education students, and less so for secondary education students. Teachers at all levels, though, employ the full range of assessment tools: paper-and-pencil tests, performance assessments (observation of and judgment about behaviors and products), and personal communication with students. Due attention

must be given to all of the techniques as well as to the different levels of learning within the cognitive, affective, and performance domains.

3. Authentic assessment (McLean, 1990), which incorporates both paper-and-pencil tests and applied performance assessment, provides a fresh view of what ought to be included in a classroom assessment design (see point 5 below). Thinking—a knowledge based, discipline oriented function—is the centerpiece of curriculum reform. All students are encouraged to think, engage in real-world problem solving, and to study curriculum that respects the integrity of the disciplines yet emphasizes the connections among them. By widening the evidential basis underlying classroom assessment, the more meaningful and relevant assessments will become in providing the basis from which to assess student achievement (see, e.g., Biggs & Collis, 1982; Glaser, 1988; Maguire, 1991). Performance on paper-and-pencil tests is alone too narrow a base for assessing student learning in the classroom.
4. Use of computers in classroom assessment is gradually becoming more common. A decade ago microcomputer software was primitive in form. But recent advances are viewed as holding significant promise for improving the capability of teachers to assess effectively in their classrooms. Even though recent research (see Hsu, 1989) suggests much more remains to be learned before microcomputer applications will become straightforward in the classroom, there still is need to discuss microcomputer advances that can now be readily incorporated in classroom assessment (e.g., in test quality control; storing and updating instructional objectives and test items and forms; keeping records and storing grades).
5. Instruction in assessment design must become a part of a meaningful classroom assessment course for preservice teachers. Planning an assessment design for a course will help integrate classroom assessment with instruction (Dassa, 1990; Nitko, 1989). Such planning organizes an overall perspective on the decisions to be made and the contexts within which they will be made. That is, the assessment design should not only describe what is to be done, but *why* it is being done and *how* it will be used. Attention to individual instruments and procedures and the trade-offs in using different instruments and strategies will then come naturally.

Conclusion

Educational assessment in Canada is not, nor will it become, extinct. Accountability testing, with its recent movement toward new assessment methodologies (e.g., in British Columbia [Bateson, personal communication, May 22, 1990] and Alberta [Horvath, personal communication, November 20, 1990]), is here to stay. And there is no question that student assessment is an essential part of teaching and learning, and that good teaching cannot exist without valid assessment of the students to be taught.

In the past, practicing teachers have either not been motivated or have not been given the opportunity to learn about assessment methodology. Assessment training must be an integral part of the preservice or undergraduate preparation of all teachers. Further, where needed, such training should be made available to teachers in the field. Adherence to the *Standards for Teacher Competencies in Educational Assessment of Students* cannot

help but improve both accountability testing and classroom assessment. Accordingly, the assessment component of the preservice education program at each Canadian university should be reviewed and evaluated in terms of what is called for in the *Standards*. Fairness and equity demand no less.

Note

1. This article was adapted from the author's President's Invited Address, Canadian Educational Research Association, June, 1990. The author is indebted to Thomas O. Maguire for helpful comments on early drafts of this paper, and to the many colleagues who willingly provided information on the measurement and assessment courses taught to preservice teacher education students at their respective universities.

References

Anderson, J.O. (1989). Evaluation of student achievement: Teacher practices and educational measurement. *Alberta Journal of Educational Research, 35*, 123-133.

Anderson, J.O., Muir, W., Bateson, D.J., Blackmore, D., & Rogers, W.T. (1990). *The impact of provincial examinations on education in British Columbia: Summary Report*. Victoria, BC: Ministry of Education.

Barringer, F. (May 13, 1990). U.S. education panel sets guidelines on assessing student performance. *New York Times*.

Bateson, D.J. (1990). Measurement and evaluation practices of British Columbia science teachers. *Alberta Journal of Educational Research, 36*, 45-51.

Benson, J. (1988). The psychometric and cognitive aspects of test-wiseness: A review of the literature. In M.H. Kean (Ed.), *Test-wiseness* (pp. 1-25). Bloomington, IN: Phi Delta Kappa.

Biggs, J.B., & Collis, K.F. (1982). *Evaluating the quality of learning: The Solo taxonomy*. New York: Academic Press.

Cole, N.S. (1988). A realist's appraisal of the prospects for unifying instruction and assessment. In *Assessment in the service of learning: Proceedings of the 1987 ETS Invitational Conference*. Princeton, NJ: Educational Testing Service.

Dassa, C. (1990). From a horizontal to a vertical method of integrating educational diagnosis with classroom assessment. *Alberta Journal of Educational Research, 36*, 35-44.

Dauncey, M. (1986). *Assessment of teachers' grading practices*. Unpublished master's thesis. University of British Columbia.

Glaser, R. (1988). Cognitive and environmental perspectives on assessing achievement. In *Assessment in the service of learning: Proceedings of the 1987 ETS Invitational Conference*. Princeton, NJ: Educational Testing Service.

Glass, G.V. (1986). Testing old, testing new: Schoolboy psychology and the allocation of intellectual resources. In B.S. Plake & J.C. Witt (Eds.), *The future of testing* (pp. 9-27). Hillsdale, NJ: Erlbaum.

Goslin, D.A. (1967). *Teachers and testing*. New York: Sage.

Gullickson, A.R. (1985). Student evaluation techniques and their relationship to grade and curriculum. *Journal of Educational Research, 79*(2), 96-100.

Gullickson, A.R. (1986). Teacher education and teacher-perceived needs in educational measurement and evaluation. *Journal of Educational Measurement, 23*, 355-368.

Gullickson, A.R. (1990). *Matching measurement instruction to classroom-based evaluation: Perceived discrepancies*. Paper presented at Burus-Nebraska Symposium on Measurement and Testing, Lincoln, NB.

Gullickson, A.R., & Ellwein, M.C. (1985). Post-hoc analysis of teacher-made tests: The goodness-of-fit between prescription and practice. *Educational Measurement: Issues and Practice, 4*, 15-18.

Gullickson, A.R., & Hopkins, K.D. (1987). The content of educational measurement instruction for preservice teachers: Professor perspectives. *Educational Measurement: Issues and Practice, 6*(3), 12-16.

Horvath, F.G. (1990). *Proposed changes to the Provincial Achievement Testing Program questionnaire results*. Edmonton, AB: Author.

Hsu, T-c. (1989). Gaps in the application of the computer to testing. *Educational Measurement: Issues and Practice, 8*(3), 4.

Impara, J.C., Divine, K., Bruce, F.A., Liverman, M., & Gay, A. (1990). *Utility of interpretive information in helping teachers' interpret standardized test score reports: The Virginia State Assessment Program*. Paper presented at the Annual Conference of the National Council on Measurement in Education, Boston, MA.

Lazur-Morris, C., Polin, L., May, R., & Barry, L. (1980). *A review of the literature on test use*. Los Angeles: University of California, Center for the Study of Evaluation. (ERIC Documentation Reproduction Service No. ED 204 411)

Linn, R.L. (1989). Current perspectives and future directions. In R.L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 1-10). New York: American Council on Education and Macmillan.

McLean, L.D. (1985). *The craft of student evaluation in Canada*. Toronto, ON: Canadian Education Association.

McLean, L.D. (1990). Time to replace the classroom test with authentic measurement. *Alberta Journal of Educational Research, 36*, 78-84.

Maguire, T.O. (1991). Evaluation of classroom instruction. In R. Short, L. Stewin, & S. McCann (Eds.), *Educational psychology: Canadian perspectives*. Toronto, ON: Copp Clark Pitman

Manke, M.P., & Loyd, B.H. (1990). *An investigation of non-achievement-related variables influencing teachers' grading practices*. Paper presented at the annual conference of the National Council on Measurement in Education, Boston, MA.

Mayo, S.T. (1964). What experts think teachers ought to know about educational measurement. *Journal of Educational Measurement, 1*, 79-86.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: American Council on Education and Macmillan.

Millman, J., Bishop, H., & Ebel, R. (1965). An analysis of test-wiseness. *Educational and Psychological Measurement, 25*, 707-726.

National Council of Measurement in Education. (1990). *Standards for teacher competence in educational assessment of students*. Washington, DC: Author.

Nickerson, R.S. (1989). New directions in educational assessment. *Educational Researcher, 18*(9), 3-7.

Nitko, A.J. (1989). Designing tests that are integrated with instruction. In R.L. Linn (Ed.), *Educational Measurement* (pp. 447-474). New York: American Council on Educational and Macmillan.

Rogers, W.T., & Bateson, D.J. (in press). The influence of test-wiseness upon performance of high school seniors on school leaving examinations. *Applied Measurement in Education*.

Rudman, H.C., Kelly, J.L., Wanous, D.S., Mehrens, W.A., Clark, C.M., & Porter, A.C. (1980). *Integrating assessment with instruction: A review (1922-1980)*. East Lansing, MI: Michigan State University.

Sarnacki, R.E. (1979). An examination of test-wiseness in the cognitive test domain. *Review of Educational Research, 49*, 252-272.

Shepard, L.A. (1989). Why we need better assessments. *Educational Leadership, 46*(7), 4-9).

Stiggins, R.J. (1990). Toward a relevant classroom assessment research agenda. *Alberta Journal of Educational Research, 36*, 92-97.

Stiggins, R.J., & Bridgeford, N.J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement, 22*(4), 271-286.

Stiggins, R.J., Conklin, N.F., & Bridgeford, N.J. (1986). Classroom assessment: A key to effective education. *Educational Measurement: Issues and Practice, 5*(2), 5-17.

Stiggins, R.J., Frisbie, D.A., & Griswold, P.A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice, 8*(2), 5-14.

Taylor, H. (1979). *Grading practices: Issues and alternatives*. Victoria, BC: Ministry of Education.

Traub, R.E. (1990). Assessment in the classroom: What is the role of research? *Alberta Journal of Educational Research, 36*, 85-91.

Wilson, R.J. (1990). Processes in evaluating student achievement. *Alberta Journal of Educational Research, 36*, 4-17.

Wilson, R.J., & Rees, R. (1990). The ecology of assessment: Evaluation in educational settings. *Canadian Journal of Education, 15*(3), 215-228.

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format

- Manuscript in triplicate.
- Typewritten or printed, double-spaced throughout.
- Abstract of approximately 100 words on a separate sheet.
- Manuscript not to exceed 6,500 words excluding graphics.
- Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.

Referencing

- Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s).
- Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format.
- All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
- Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes*; any citations in notes follow the same format as other references.

Graphics

- Each table and figure on a separate page at the end of the manuscript.
- Figures, charts, and diagrams in camera-ready form.
- Number tables in Arabic numerals and center over each heading, e.g., *Table 1.*
- Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1: ...*
- Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*

Style

- *AJER*'s editorial style conforms closely to the *Publication Manual of the American Psychological Association.*
- Editorial changes may be made to manuscripts.
- For spelling, consult *Webster's New Collegiate Dictionary;* spelling in quoted material remains as in the original.

Diskettes

If manuscript is accepted for publication and has been prepared using a Macintosh, IBM, or IBM compatible microcomputer, please send diskette. Microsoft Word, MultiMate, WordPerfect, WordStar, Xerox Writer, XyWrite, and DisplayWrite 3 **only** are acceptable.
Important: Please indicate program used.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index*; appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*

ajer

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXVII NUMBER 3 SEPTEMBER 1991

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Thomas O. Maguire

Consulting Editors

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

Janet Donald
McGill University

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Lilian Katz
University of Illinois

Les McLean
Ontario Institute for Studies in Education

John McLeod
University of Saskatchewan

Linda Phillips
Memorial University of Newfoundland

Lauran Sandals
University of Calgary

Keith Sullivan
Dalhousie University

Phil Winne
Simon Fraser University

Ruth Wright
University of Ottawa

Faculty Advisory Committee

Harvey Zingle (Chair)
Fern Snart
Erwin Miklos
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Second Class Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXVII, Number 3, September 1991

Special Section

W. TODD ROGERS 193 Introduction to Special Section on Gender Differences in Achievement

SHARON HAGGERTY 195 Gender and School Science: Achievement and Participation in Canada

JEFFREY W. BULCOCK, MORLEY E. WHITT, MONA J. BEEBE 209 Gender Differences, Student Well-Being and High School Achievement

GAALEN ERICKSON, SANDRA FARKAS 225 Prior Experience and Gender Differences in Science Achievement

BIKKAR S. RANDHAWA 241 Gender Differences in Academic Achievement: A Closer Look at Mathematics

Perspectives

JOHN HATTIE 259 The Burt Controversy: An Essay Review of Hearnshaw's and Joynson's Biographies of Sir Cyril Burt

WILLIAM J. HAGUE 277 Kohlberg's Legacy— More than Ideas An Essay Review

ISSN 0002-4805

The Alberta Journal of Educational Research *Vol. XXXVII, No. 3, September 1991, 193-194*

Introduction to Special Section on Gender Differences in Achievement

W. TODD ROGERS
Guest Editor

It is hard to find a more controversial topic in recent behavioral and social science research than that of gender differences in achievement. During the last 30 years or so, a great deal of research has been devoted to explaining these differences. Yet debate persists as to the extent of and reasons for such differences and the validity and significance of the findings and explanations provided.

Much of the research has been conducted in the areas of science and mathematics, likely because of the prominence given to mathematics and science as cornerstones for effective and meaningful economic development and international competitiveness in today's world. Notably, and of concern, is the observation that women are underrepresented, some would argue dramatically, on university science and mathematics faculties and in scientific and technical careers. For many, the reasons for this underrepresentation are traceable to the failure to account for differential performance in school, particularly at the secondary school level, where students begin to select the program of studies they will follow and the science and mathematics courses in which they will enroll, and where there is general agreement that gender differences in achievement begin to emerge.

The four articles included as a special section on gender differences in this issue of *AJER* address the issue of gender differences in science and mathematics during the school years in a Canadian context. Collectively, the articles convey an excitement that research may actually make a difference, that research findings can be used to develop relevant and appropriate procedures for meeting the learning needs of both male and female students in science and mathematics at all levels of schooling. In agreement with research conducted in the United States and elsewhere, there is consistent reference that gender differences in science and mathematics begin to emerge about the onset of adolescence. We must be doing something to young adolescent girls and boys and/or we are failing to take proper account of prior learnings and experiences girls and boys enjoy to socialize them into different patterns of behavior, choice, opportunity, and expectations.

Individually, each article contributes knowledge leading to a greater understanding of gender differences in science and mathematics, and each raises questions that underscore the need for additional research. In the first article, Sharon Haggerty comprehensively surveys the Canadian scene in science and introduces many of the themes that appear in the subsequent articles. Jeff Bulcock, Morley Whitt, and Mona Beebe then address the question "Why are female high school students in Newfoundland outperforming boys in all school subjects except physics?" Rejecting both biological-

ly based theory and culturally based theory, they test a demographic model for accounting for the gender differences in school achievement they observed using causal modeling. In the third article, Gaalen Erickson and Sandra Farkas address the issue of gender differences from a qualitative perspective. They explore the apparent impact of prior experience, in and out of school, on student performance on multiple-choice science achievement items. In particular, they address the question "Why do girls have less success 'applying' concepts to practice than do boys?" Employing the notion of effect size and advanced data analysis procedures, Bikkar Randhawa examines possible reasons for differential effects in mathematics at three different grade (age) levels. To clarify his findings, he simultaneously investigates gender differences in reading using the same analytical procedures.

One disturbing trend seems to be that while researchers are aware of and trying to do something about the issue of gender differences, policy makers and school officials, particularly teachers, may be less convinced. There may be better opportunities and riches for all if the findings of these studies are carefully considered and used by those who work with school students.

A possible disappointment with this special section of *AJER* on gender differences is that the articles, taken as a set, emphasize science and mathematics. Little attention is given to other subject areas such as literature and the fine arts. Three of the four articles were presented in separate sessions of the 1990 Annual Conference of the Canadian Society for the Study of Education/Societé Canadienne pour l'Étude d'Éducation. All the presenters were asked if they wished to have their presentation considered for publication in *AJER* with the understanding that each article would undergo the regular review process. All agreed. Each was asked to nominate one or two other researchers whom he or she knew were working in the area of gender differences. Of the nominees, all were working in the areas of science and mathematics. Does the failure to attract researchers investigating gender differences in other subject areas reflect the absence of such research? While, as mentioned earlier, science and mathematics are key subjects, so are other subjects in the humanities and social science areas. A second question to ask and which arises from the findings presented here and elsewhere is not "Do gender differences occur?" but "Under what circumstances do differences exist and what can be done to provide opportunities for both girls and boys to have their understandings reflected in achievement?" There is a task here for test-makers. We invite you to analyze the findings of the four articles presented here and the literature as we, with our interests in testing, have just done to identify other questions that can be used to direct and clarify the nature and causes of gender differences observed in achievement not only in science and mathematics, but also in the humanities and social sciences.

In closing, we wish to thank the reviewers for their insightful comments and constructive suggestions. The set of articles is stronger as a result of their contributions and concern for gaining greater understanding of gender differences and ways they may be positively addressed in school.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 3, September 1991, 195-208*

SHARON M. HAGGERTY
University of Western Ontario

Gender and School Science: Achievement and Participation in Canada

A review of recent studies of gender differences in science achievement and participation suggests that achievement differences cannot be understood or addressed without considering the effects of both attitude and participation. Canadian findings on achievement and participation support the suggestion that females and males approach school science differently and that those approaches are an important factor in the achievement differences that are observed. Enrollment data from ministries and departments of education reveal marked differences in participation in elective sciences, with the smaller provinces appearing to be more successful at encouraging students to persist in their science studies at the senior high school level.

The issue of gender differences in science and mathematics achievement has probably received more attention than in any other subject. Concern for girls and science has existed internationally since the publication of the results of the first international science assessment (Comber & Keeves, 1973; Kelly, 1978). Arising from this concern, an international organization, Gender and Science and Technology (GASAT), was established in 1981 and continues to meet biennially[1] to work toward achieving educational equity in science, mathematics, and technology. Canadians have participated in GASAT since its inception.

The Science Council of Canada (1984) conducted a major study of science education in Canada. Among its many recommendations were two specifically developed to increase the participation of girls in science. At the provincial level:

> Ministries should outline, in all policy documents relating to science and technology education, concrete steps to be taken by educators to increase the participation of girls in science and technology. (p. 48)

and at the national level:

> The Minister of State for Science and Technology and the Minister responsible for the Status of Women should together sponsor a large-scale public information program directed especially towards parents and designed to raise public awareness of the need for girls and young women to participate in science and technology education. (p. 63)

Attention has also been drawn to gender differences in favor of males in science achievement[2] as measured in various provincial (e.g., Bateson,

Sharon Haggerty is an assistant professor of curriculum studies at the University of Western Ontario. Her current research interests are related to perceptions of the nature of science and science teacher education.

Anderson, Dale, McConnell, & Rutherford, 1986) and international assessments (e.g., Connelly, 1987; Connelly, Crocker, & Kass, 1985; Lapointe, Mead, & Phillips, 1989).

While gender differences continue to be reported, debate persists as to their extent and their validity and significance. Not all studies have detected achievement differences; and in studies where differences are found there are a number of views as to the reasons why they occur. Some suggest that gender differences can be accounted for by differential course taking, while others believe the causes are more deeply embedded in factors related to differential socialization of females and males.

As well as achievement and participation in science, attitudes toward science are considered in this article. A recent longitudinal study of influences on attitude toward and achievement in high school science conducted in the United States revealed a strong relationship between achievement and attitude (Simpson & Oliver, 1990). It was found that school variables strongly influenced attitudes, and that attitudes of grade 10 students successfully predicted which students would not pursue future elective science courses in 90% of the cases studied. Hence the view taken in this article is that achievement and attitudes should not be examined in isolation from each other. It is also suggested that merely examining gender differences is little more than an academic exercise. We must also be prepared to develop and implement strategies aimed at ensuring that both females and males in Canada are equally successful in school science. Suggestions as to how that might be addressed are presented.

Achievement in Science

As indicated above, recent science assessments in Canada and elsewhere have found gender differences in favor of males in science achievement, and that these differences first became noticeable at early adolescence, increasing during the high school years. Individual science teachers, however, may dispute their existence, stating there are no differences between the achievement of females and males in their classes (Sadker & Sadker, 1986; Tobin & Garnett, 1987). Moreover, average grade 12 final marks obtained from the British Columbia Ministry of Education reveal little difference between the final marks of male and female grade 12 students in science in that province (Kozlow, personal communication, 1986). Thus there is conflicting evidence as to whether gender differences in science achievement really do exist.

In order to account for this apparent discrepancy, Haggerty (1987) suggested that there may be two distinctly different kinds of achievement measures in science. In the first instance, a test score is being considered. Generally, the test is externally prepared and presented in a multiple choice format. It covers a wide range of content that the students are expected to have encountered in their previous science studies. The students generally do not know what specific content is covered on the test. Further, the context of the items may be quite different than the context in which the particular scientific principles were studied in school, and at times the students do not have the opportunity to study a particular body of content to be tested. Perhaps most importantly the students are aware that the results of such

tests will not influence their marks in any way. The second type of achievement measure is a mark or grade obtained by students for a particular science course in school. Such a mark is a composite of test scores and marks/grades received on various assignments. The tests in this instance are very different in nature. Generally, the content tested has been recently studied in a particular science course. The students study content identified by the teacher. A portion of the mark may be given for work effort in class and for completion of assigned homework. Haggerty suggested these two kinds of achievement measures (i.e., scores on standardized tests and course marks) are actually measuring distinctly different sorts of achievement. To investigate this possibility, she examined data from three British Columbia sources.

In the first analysis, scores on the British Columbia science assessment achievement scales showed superior performance by males for almost all science topics (Bateson et al., 1986; Hobbs, Boldt, Erickson, Quelch, & Sieben, 1979; Taylor, Hunt, Sheppy, & Stronck, 1982). The gender differences were greatest for physical and earth science topics and on questions requiring higher level thinking skills. Moreover, the differences increased with age. In the second analysis, final marks obtained by grade 12 students in British Columbia were examined. In contrast to the assessment results, these final marks are comprised of two equally weighted components: one the mark assigned by the classroom teacher and the other the score on a provincial school leaving examination.[3] It was found that in the sciences the average marks obtained by females ranged from 0.8% less than those obtained by males in chemistry to 2.9% more than males in physics. Thus looking at gender differences on these two kinds of measures (i.e., provincial science assessment achievement scales and blended final marks in biology, chemistry, and physics) yielded different findings.

In the third analysis, Haggerty (1986, 1987) drew on data from a study of a grade 9 science class during a unit on heat and temperature. In the study, scores were obtained on a posttest measure that tested how well students could apply scientific knowledge related to a variety of phenomena[4] they had studied in school. The contexts of the items in this posttest, and in a similar pretest administered prior to instruction, were different than those used by the teacher to teach the unit. And, as in the case of the provincial assessment, the students could not study for this test by learning a particular set of facts and the students knew the results of the test would not be included in their course marks. The pretest revealed no significant gender differences in understanding the heat and temperature concepts tested. Similarly, there were no significant gender differences in the scores obtained by the students on the unit test (set by the teacher and limited to questions that could be answered correctly by a student who had memorized key definitions and other facts), and on combined notebook, homework, and unit test marks assigned by the teacher for the unit as a whole. In contrast, the posttest did reveal significant gender differences. After completing the heat and temperature unit, the boys seemed better able than the girls to apply what they had learned to the novel situations presented in the posttest measure. This finding suggests that males understood the science they learned more effec-

tively than the females, in spite of the marks obtained in their science course. Similar findings have been reported by Bejar and Blew (1981), De Boer (1984), Levin and Fowler (1984), Maccoby and Jacklin (1974), Sadker, Sadker, and Klein (1986), Simpson and Oliver (1990) and Zarega, Haertel, Tsai, and Walberg (1986). Researchers have attempted to identify plausible explanations for why girls receive course marks comparable to those of boys, yet apparently are not able to apply what they learned as well as boys. One suggestion is that better work habits of females provide them with an advantage on school achievement measures (De Boer, 1986; Maccoby & Jacklin, 1974; Walberg, 1969; Zarega et al., 1986). Ridley and Novak (1983) have suggested that females are socialized toward rote learning strategies. Concern for right answers among females is often noted by teachers and was clearly evident in the classroom study conducted by Haggerty (1986). Such findings reinforce the contention that the important question is not whether gender differences exist, but rather, under what circumstances do gender differences exist and, most importantly, what can be done to provide both females and males with opportunities to understand the science they learn and to have that understanding reflected in their achievement?

As noted earlier, there appear to be three plausible explanations as to why males outperform females on standardized tests (such as the British Columbia Science Assessment and the American National Assessment of Educational Progress or NAEP). Eylon and Linn (1989) are among those who support the differential course taking hypothesis, suggesting that because males take more science, they perform better. Erickson and Erickson (1984), however, using data from the British Columbia science assessment, dispute that. Similarly, Mullis and Jenkins (1988), using American NAEP data, found that course enrollment had no effect on the magnitude of the gender difference. In the second international science assessment (International Association for the Evaluation of Educational Achievement, 1988), when students enrolled in their final year of secondary science were tested by subject, there were still significant gender differences in Canada and most other countries (differences for chemistry and physics were greater than for biology).

In their recent review of gender and mathematics and science, Linn and Hyde (1989) conclude that informal as well as formal learning experiences were important. They noted:

> Gender differences in science knowledge and scientific reasoning reflect both formal and informal learning experiences. Overall gender differences in science (a) are larger for science knowledge than for science processes, (b) are declining among high school students, and (c) accompany gender differences in formal and informal learning experiences. (p. 22)

Other British Columbia work by Erickson and Farkas (1987) suggest that males and females tend to draw on different sources of personal knowledge when confronted with questions on a standardized test, and that a lack of relevant prior informal experiences contributes to females' inferior performance. Students were questioned as to what they had been thinking about as they chose their answers to science assessment questions on which sig-

nificant achievement differences between females and males had occurred. Females more than males indicated that they recalled an incident that had occurred in school science classes, whereas males more than females recalled an incident outside of formal school science. Similarly, Linn and Hyde noted that "males compared with females report substantially more informal experience with physical science" (1989, p. 22). Haggerty (1987) has suggested that females may do better on test questions that can be answered by rote learning, whereas the reverse may be true for males. This suggestion is consistent with the findings of both Erickson and Farkas (1987) and Linn and Hyde (1989) and could also account for males' superior performance on higher level assessment questions and for females' superior performance on lower level questions. Similarly, school science course marks are more likely to reflect good work habits and to be based on tests that ask predominantly lower level questions, thereby contributing to the relatively higher achievement of females on these measures.

One additional explanation has been proposed to account for the differences on British Columbia assessment scales. Bateson and Parsons-Chatman (1989) suggest that the differences may be due to gender bias in the content of the items. Erickson and Farkas's (1987) findings do not support that explanation. At best, the content of the items appears to be insufficient to account for the explanations given by students in Erickson and Farkas' study.

The view that any one factor such as differential course taking, prior informal experiences, or content of items is responsible for gender differences in achievement is not supported by the studies reviewed above. There is now considerable evidence that achievement is influenced by a complex of factors.

Attitudes Toward Science

In a meta-analytic study of gender and science, Steinkamp and Maehr (1984) concluded that although in general females were more motivated to achieve than males, the reverse was true at the junior high level in the United States. There, adolescent males' motivation to succeed exceeded that of females and the gap between the two had been increasing during the preceding decade. Since then, Simpson and Oliver (1990) have conducted a longitudinal study with adolescents in the United States. Using extensively revised existing instruments, as well as developing additional instruments for their work, they finally produced and utilized a 60-item, 14-subscale attitude survey. In their study, females scored consistently higher on motivation to achieve in science (among other subscales), whereas males had significantly more positive attitudes toward science. Simpson and Oliver noted, as have others (e.g., Duncan & Haggerty, 1985; Yager & Yager, 1985), a general decline in attitudes toward science during the adolescent years. In a Saskatchewan study of 43 grade 7 to 9 science classes, Wideen (1974) found that interest in science was greater for males than females, and that for students who indicated that a subject other than science was their favorite subject, there was a decline in science interest from grade 7 to grade 9. The decline did not occur among students who chose science as their favorite subject in school.

Similarly, Kelly (1986) included gender as a factor in a longitudinal study of science attitudes. She found that from ages 11 to 13 children's attitudes toward science declined significantly and noted that "their personal liking for science showed a particularly marked drop, especially among girls" (p. 402). These declines in adolescents' attitudes toward science should be of concern to those involved in science education during the middle school/junior high school years.

Linn and Hyde (1989) discussed three aspects of attitude (or what they call *psychosocial*) differences: aggression (e.g., males' greater participation in class activities and discussions), confidence (e.g., males overestimate their abilities, females are realistic), and interest. They noted that gender differences were heterogeneous and interacted both with the particular situation and with cultural factors. Linn and Hyde concluded that the only interest factor that seemed to influence persistence in science and mathematics was perceived usefulness, and suggested that this may be related to the frequently reported concern that the science taught in schools is not seen as addressing relevant topics or problems. As most provinces have expressly developed curricula with relevance to students (particularly at the middle school/junior high school level), this is a concern. It appears that although curriculum developers believe the curriculum addresses topics that are socially relevant to adolescents, that relevance is not apparent to the students. If Linn and Hyde are correct that science participation is related to perceived relevance, then perhaps we might hope to increase participation if we can make that relevance apparent to the students. Similarly, we might look to perceived relevance as we ponder explanations for low participation in elective science courses.

Simpson and Oliver (1990) have come to somewhat different conclusions. Data were initially collected on 4,500 grade 10 students. To determine the extent to which further participation in science could be predicted, additional data were collected five years later from these students. Student attitudes toward science and success in science in grade 10 were a strong predictor of subsequent science achievement in high school, and science self concept at grade 10 was a significant predictor of future course participation. In particular, students with low science attitudes did not take elective science courses. The authors concluded, "A major finding of this study, therefore, is that attitudes toward science play a key role in influencing the amount of exposure to science a student experiences" (p. 13). The longitudinal nature of this study appears to have allowed more precise identification of the relationship between attitude and both achievement and participation in elective sciences. Simpson and Oliver also found, contrary to their expectations, that school variables were more strongly related to attitude toward science than were family variables. This is encouraging for educators who are sometimes discouraged from attempting to address student attitudes due to the perception that the home is the major influence on student attitudes. Simpson and Oliver's finding suggests that it is important to consider the school environment when addressing attitude issues.

Participation in Science

Two aspects of participation are considered—choosing to enroll in elective science courses and student participation in discussions and activities in the science classroom.

Enrollments in Elective Science

Studies reviewed in the previous section suggest that attitude is an important factor in a student's choice to enroll in science courses. Ridley and Novak (1983) have suggested that the socialization of females toward rote learning strategies may be related to low female participation in advanced science and mathematics courses. In order to examine the pattern of student enrollments in science, representatives of each of the 10 provincial departments/ministries of education were asked to provide student enrollment data for grade 10 to 12 science courses. In the first analyses of these data, male and female enrollment in grade 11 sciences were compared with the corresponding populations in each province. Generally, grade 11 marks the point at which enrollment in a science course becomes optional rather than compulsory, that is, science becomes an elective rather than a required course. To obtain an estimate of the total number of females and males who might potentially enroll in a grade 11 science, the population of 15- to 19-year-olds (Statistics Canada, 1988) was divided by five. There is greater participation by both males and females in some provinces than in others. For example, in Saskatchewan the percentages of females enrolled in biology, chemistry, and physics were respectively, 76%, 61%, and 47%. The corresponding data for Ontario were 43%, 32%, and 20%. This result illustrates a second point: higher enrollments in grade 11 science tend to be inversely related to total grade 11 population size across the six provinces for which there were sufficient data (see Table 1). Thirdly, examination of the corresponding

Table 1
Estimated Participation of Females and Males in Grade 11 Biology, Chemistry, and Physics by Province
1987/1988

Province	Female Pop'n X1000 (15-19)	Females			Male Pop'n X1000 (15-19)	Males		
		Biol.	Chem.	Phys.		Biol.	Chem.	Phys.
B.C.	101.8	45%	34%	19%	106.9	40%	34%	34%
Alberta	89.3	67	58	24	92.9	54	55	38
Saskatchewan	38.3	76	60	47	40.3	64	55	50
Ontario	336.4	43	31	20	352.7	29	34	28
N.B.	29.4	34	40	28	31.0	30	36	33
P.E.I.	5.2	73	60	30	5.6	49	46	39

Note: Data not included for some provinces because of insufficiency (Manitoba and Newfoundland/Labrador) or because not all sciences are offered at the grade 11 level (Nova Scotia and Quebec).

course percentages for females and males reveals that across all provinces a greater percentage of females than males enroll in biology; in contrast, the percentage of females in physics is lower. The pattern of differences for chemistry is less clear.

To further examine the relative participation of females and males enrolled in each science course offered in grades 10, 11, and 12 in each province, the proportion of males and females enrolled in each subject is shown in Table 2. A number of observations can be made. First, in agreement with the data provided in Table 1, more males than females study biology, and, except among francophones in Ontario and Quebec, more males than females studied physics. For chemistry, the proportions of males and females were more nearly equal. There were more males in chemistry in British Columbia and Newfoundland and among central Canadian anglophones, but more females among the francophones and in the maritime provinces. The greater participation by females in francophone classes is interesting and probably warrants further examination.

Although it is impossible to explain these course enrollment variations with certainty, some patterns have been identified and some speculations can be made. The smaller, more rural regions of Canada (i.e., the prairies and Atlantic provinces) tend to have higher science participation. This is particularly so for Saskatchewan and for females in Prince Edward Island. New Brunswick, however, is an exception. New Brunswick differs from the others in that it offers a greater variety of academic and nonacademic courses. In New Brunswick, students who will be continuing to postsecondary education take different courses than those who will not. New Brunswick also has a relatively higher proportion of females in physics than any other province except Saskatchewan (not including francophone classes) and has made an effort to attain gender equity in education (Department of Education, 1987). Physics participation by females is particularly dismal in the final year of schooling in British Columbia, Ontario, and Alberta. These three provinces are more industrialized and economically advantaged than the others.

The differences between Saskatchewan and Alberta are of interest. In both of these provinces approximately half of those enrolled in chemistry are female. In Alberta, however, the typical gender bias (more females take biology, more males take physics) is more pronounced in biology and physics. Historically, these provinces have much in common. There are differences in the dominant economic base and the dominant political philosophy of these neighboring provinces. In Alberta the oil industry is particularly significant in the provincial economy, whereas farming dominates in Saskatchewan.

Overall, among anglophones the gender bias is most pronounced in the three provinces that not only have the larger populations, but also are more industrialized than the others (i.e., Alberta, British Columbia, and Ontario)—the question is, is this merely a coincidence, or is there any possible significance to such trends?

Differences in course participation across Canada illustrate the complexities involved in this issue. Sociocultural influences are undoubtedly involved. However, it behooves those of us in the educational community to focus on school and curriculum variables in the hope that we can have some

Table 2
Percentage of Females in Canadian Secondary Science Classes, 1987/88

Province	Science	Biology			Chemistry			Physics		
Grade:	10[a]	11	12	13	11	12	13	11	12	13
B.C.	50	57	61		49	43		35	21	
Alberta	53/50/49[b]	55	59		50	51		38	33	
Saskatchewan	50	53	55[c]		51	52		47	44	
Manitoba[d]			(57)			(49)			(36)	
Ontario Fr.[e]	56	63	50	63	53	58	47	53	52	41
Eng.	51	59	61	59	48	48	44	43	39	32
Quebec Fr.	53/50[f]				53					
Eng.	49/38				45		47			
N.S.	50/46[g]		58		52	52		45	36	
P.E.I.	54	60	61		55	56		42	43	
N.B: Fr.		68[h]	63		53	57		48	47	
Eng.		53	58		51	51		45	44	
Nfld/Lab.[i]		60			43			27		

[a]In most provinces general science is compulsory in grade 10.

[b]In Alberta in 1987/88 each of biology, chemistry, and physics was offered in grade 10. Each student was required to take at least *one* of the three courses. In 1990/1991 general science will replace the three separate courses.

[c]Francophone biology 12 classes were 57% female.

[d]Data for Manitoba were estimated from a bar graph; grade was not indicated.

[e]Ontario, Quebec, and New Brunswick maintain separate records for francophone and anglophone classes. Ontario is the only province that provides a 13th year of schooling.

[f]In Quebec there is no grade 12. Chemistry and physics are offered at years 10 and 11, respectively. Biology is not offered after year 9.

[g]In Nova Scotia biology and physical science are offered in grade 10.

[h]The first year of biology in New Brunswick is a grade 10 course.

[i]Newfoundland/Labrador did not respond to the request for data. Figures given are taken from Scott (1982) and are for the final year of secondary school which at that time was grade 11. Since then grade 12 has been introduced as well.

impact on how students experience school science. The following section examines how the school science experience differs for females and males.

Participation in Science Classes

Differences in the way females and males interact and respond in the classroom have been well documented in a variety of subject areas and at all age levels (Becker, 1981; Bossert, 1981; Haggerty, 1987; Linn & Hyde, 1989; Sadker & Sadker, 1986; Tobin & Gallagher, 1987; Wienekamp, Jansen, Fickenfrerichs, & Peper, 1987; Whyte, 1984). For example, females appear less likely to engage in high risk situations (Lolley, Davies, & Scott-Hodgetts, 1986). And, although females are more likely to respond "I don't know," they are less likely to give wrong answers than males (Linn, De Benedictis, Delucchi, Harris, & Stage, 1987).

In Haggerty's (1986) classroom study of a grade 9 science class studying heat and temperature, the frequency of different forms of teacher-student verbal interactions was recorded during two weeks of class discussion (Table 3). Although only 39% of the students in the class were male (9 of 23), males were called upon to answer 58% of the questions and males initiated 67% of the student-initiated dialogue (see Table 3). When students did initiate dialogue, females were more likely to provide information, whereas males were more likely to ask a question. Although females comprised 61% of the class, they responded to only 42% of the teacher's questions. However, 75% of their responses were accepted as correct by the teacher. The males, who responded to 58% of the questions asked by the teacher, provided correct responses 64% of the time. When the teacher responded to students' responses, she was more likely to encourage further responses or to provide explanations when responding to a female student, and less likely to redirect a question or ignore a response from a female student. Taken together, these two forms of response (student, teacher) suggest that in spite of the teacher's positive and encouraging responses to female students, females were still much less responsive than the males. These findings are consistent with those reported in other studies in the United States and Australia (Sadker & Sadker, 1986; Tobin & Gallagher, 1987).

Kahle, Matyas, and Cho (1985), Tobin and Garnett (1987), and Zimmerer and Bennett (1987) all found that males reported different perceptions of their activities in school science than did females in the same classes. Males reported greater time spent in activities than did females in the same classes. Haggerty (1987) suggested these differing perceptions may be related to the ways in which males and females go about doing science. Males appear to be more involved in manipulative laboratory activities and females to be more concerned with finding the right answers.

The issue of male dominance in classrooms is a difficult one. In classes taught by teachers who claimed to actively encourage participation of female students, males still dominated the activities and discussions, a situation that was not recognized by the teachers concerned (Duncan & Haggerty, 1985; Sadker & Sadker, 1986; Tobin & Garnett, 1987). Observers noted that even when a teacher deliberately called on students proportionally, males were more likely to attempt a response when they were unsure about a question. This provided an advantage for the males, as even an incorrect answer provides the teacher with a basis for continuing to probe and draw out a more acceptable response. Alternatively, a response of "I don't know" gives the teacher nothing to build on and she or he typically passes on to another student. If we wish to overcome this problem, teachers will need to develop strategies to draw out and encourage all students, both female and male, to make an attempt to answer questions. Teachers must convince students (and be convinced themselves) that giving the correct answer should not be the major goal in science classrooms. Rather, the process of developing and evaluating alternative answers is the essence of doing science. As Tobin and Garnett (1987) have noted, the first step is for teachers to recognize the tendency for students to want to identify a single correct answer to any question as problematic.

Table 3
Student-Teacher Dialogue During Class Discussion

	Male		Female		Total
	N	%	N	%	N
Number of Students	9	39.1	14	60.9	23
Students Called on by Teacher	43	58.1	31	41.9	74
Student Responses to Questions					
Correct Answer	100	63.7	86	75.4	186
Incorrect Answer	14	8.9	13	11.4	27
Partially Correct	13	8.3	10	8.8	23
Alternative Belief	21	13.4	6	5.3	27
No Response	6	3.8	2	1.8	8
Total	157	100.0	114	100.0	271
Student Initiated Dialogue					
Question	37	55.2	12	36.4	49
Information	30	44.8	30	60.6	50
Total	67	100.0	32	100.0	99
Teacher Responses to Students					
Encourage, Explore	100	33.9	71	38.2	171
Acknowledge Answer	65	22.0	44	23.7	109
Wrong Answer	11	3.7	2	1.1	13
Provide Information	38	12.9	19	10.2	57
Explanation	28	9.5	27	14.5	55
Demonstration	2	0.7	0		2
Redirect Question	7	2.4	1	0.5	8
Repeat	24	8.1	19	10.2	43
Ignore or Dismiss	16	5.4	2	1.1	18
Managerial	4	1.4	1	0.5	5
Total	295	100.0	186	100.0	481

Conclusions

This article reviews recent studies of gender differences in science achievement, attitude, and participation. Achievement differences are often considered to be related to attitudes toward science and to different participation rates in science. The studies reviewed suggest that these three factors are interrelated and that each contributes to the other.

At the present time, although participation varies across Canada, generally fewer than half of our students choose to study elective science courses. Differences in participation of francophone and anglophone females were unexpected. Some sociocultural factors that might possibly influence participation have been speculated on, but these would require considerably more investigation before any claims could be made.

Studies completed in Canada and elsewhere are reviewed in this article and suggest that for females more than for males success in school science is attained by learning right answers. It is suggested that the result of this approach to learning is that girls more than boys frequently do not understand the science they learn. This approach to learning may contribute to females' lower level of achievement on standardized tests. In turn, failure to understand the science that is being learned leads females to feel that science, particularly physics, is not for them. Consequently, they are less likely to participate in elective physical science courses. If females' disadvantage with respect to both participation and achievement is to be overcome, it will be necessary for middle school/junior high school science to be taught in a more meaningful and relevant way, involving girls as well as boys as active participants in science classes. Oliver and Simpson (1988) stress the importance of "structuring the science classroom learning environment to improve the student's belief that he or she can succeed in science" (p. 154). Parker and Rennie (1986) found that teacher in-service aimed at reducing gender stereotyping of science had positive effects on 10-year-olds in Australia. Kelly (1986) reported that interventions aimed at improving attitudes of 11- to 13-year-olds in several British schools also had positive effects. Duncan and Haggerty (1985) found that in one of two Canadian high schools they studied, the science department head/chemistry teacher successfully recruited large numbers of female students into grade 11 chemistry. Informal anecdotal information from teachers also suggests that individual, enthusiastic teachers can make a difference. Studies currently underway in British Columbia are aimed at identifying characteristics of schools that do have above average participation of females in science. Hopefully, this work will assist teachers and schools to identify ways of increasing participation in science.

Overall, the findings from these studies are encouraging, and teachers should feel empowered by such research. Teachers can have an impact on achievement and participation of their students in school science.

Acknowledgment

The author would like to thank Phil Nagy, the anonymous reviewers, and particularly Todd Rogers for their many helpful suggestions.

Notes

1. The 1993 GASAT conference will be held in Canada at the University of Waterloo. Only once before has GASAT met in North America (at the University of Michigan in 1987).
2. Although this article specifically deals with science, many of the situations and views described apply equally to mathematics.
3. At the time of this study, grade 12 final marks in academic courses were determined by a combination of a teacher-assigned mark and the score obtained on the provincial examination. The two scores were equally weighted.
4. For example, students were told that if there were a large ice cube and a small ice cube in a beaker of water, the large ice cube would take longer to melt than the smaller one. They were then asked if that meant it took more heat to melt the larger cube. On both the pretest and the posttest many students replied no, because ice always melts at 0 degrees. Such students clearly did not distinguish heat and temperature appropriately, although that was a major goal of the unit, and all students were able to define the two terms correctly on the unit test.

References

Bateson, D., Anderson, J., Dale, T. McConnell, V., & Rutherford, C. (1986). *British Columbia science assessment: General report*. Victoria, BC: Ministry of Education.

Bateson, D.J., & Parsons-Chatman, S. (1989). Sex-related differences in science achievement. *International Journal of Science Education, 11*, 371-385.

Becker, J.R. (1981). Differential treatment of females and males in mathematics classes. *Journal for Research in Mathematics Education, 12*, 40-53.

Bejar, I.I., & Blew, E.O. (1981). Grade inflation and the validity of the Scholastic Aptitude Test. *American Educational Research Journal, 18*, 143-156.

Bossert, S.T. (1981). Understanding sex differences in children's classroom experiences. *Elementary School Journal, 81*, 255-266.

Comber, L.C., & Keeves, J.P. (1973). *Science education in nineteen countries*. Stockholm: Almquist and Wiksell.

Connelly, F.M. (1987). *Ontario science education report card: Canadian national comparisons*. Toronto, ON: Queen's Printer.

Connelly, F.M., Crocker, R.K., & Kass, H. (1985). *Science education in Canada: Policies, practices and perceptions*. Toronto, ON: Ontario Institute for Studies in Education.

De Boer, G.E. (1984). A study of gender effects in the science and math course-taking behavior of a group of students who graduated from college in the late 1970's. *Journal of Research in Science Teaching, 21*, 95-103.

De Boer, G.E. (1986, April). *Perceived science ability and effort among males and females: Relationships to persistence in undergraduate science studies*. Paper presented to the Annual Meeting of the American Educational Research Association, San Francisco.

Department of Education. (1987). *Educational opportunities for girls and boys: A commitment to sex-equity*. Fredericton, NB: Author.

Duncan, M.J., & Haggerty, S.M. (1985). *Action plan for girls in science in British Columbia*. Report prepared for the Department of the Secretary of State, Canada, Women's Programs.

Erickson, G.L., & Erickson, L.J. (1984). Females and science achievement: Evidence, explanations, and implications. *Science Education, 68*, 63-89.

Erickson, G., & Farkas, S. (1987). Prior experience: A factor which may contribute to male dominance in science. In J.S. Daniels & J.B. Kahle (Eds.), *Contributions to the Fourth GASAT Conference* (vol. 1, pp. 49-56), Ann Arbor, MI: University of Michigan.

Eylon, B., & Linn, M.L. (1989). Learning and instruction: An examination of four research perspectives in science education. *Review of Educational Research, 58*, 251-301.

Haggerty, S.M. (1986). *Learning about heat and temperature: A study of a grade 9 science class*. Unpublished doctoral dissertation, University of British Columbia.

Haggerty, S.M. (1987). Gender and science achievement: A case study. *International Journal of Science Education, 9*, 271-279.

Hobbs, E.D., Boldt, W., Erickson, G., Quelch, T., & Sieben, G. (1979). *British Columbia science assessment: General report*. Victoria, BC: Ministry of Education.

International Association for the Evaluation of Educational Achievement. (1988). *Science achievement in seventeen countries*. Oxford: Pergamon.

Kahle, J.B., Matyas, M.L., & Cho, H. (1985). An assessment of the impact of science experiences on the career choices of male and female biology students. *Journal of Research in Science Teaching, 22*, 385-394.

Kelly, A. (1978). *Girls and science*. Stockholm: Almquist and Wiksell.

Kelly, A. (1986). The development of girls' and boys' attitudes to science: A longitudinal study. *European Journal of Science Education, 8*, 399-412.

Lapointe, A.E., Mead, N.A., & Phillips, G.W. (1989). *A world of differences: An international assessment of mathematics and science*. Princeton, NJ: Educational Testing Service.

Levin, J., & Fowler, H.S. (1984). Sex, grade, and course differences in attitudes that are related to cognitive performance in secondary science. *Journal of Research in Science Teaching, 21*, 151-166.

Linn, M.C., De Benedictis, T., Delucchi, K., Harris, A., & Stage, E. (1987). Gender differences in National Assessment of Education Progress items: What does "I don't know" really mean? *Journal of Research in Science Teaching, 24*, 267-278.

Linn, M.C., & Hyde, J.S. (1989). Gender, mathematics, and science. *Educational Researcher, 18*(8), 17-19, 22-27.

Lolley, M., Davies, S., & Scott-Hodgetts, R. (1986). Teachers as researchers. *Mathematics Teaching, 118*, 46-47.

Maccoby, E.E., & Jacklin, C.N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Mullis, I.V.S., & Jenkins, L.B. (1988). *The science report card, elements of risk and recovery: Trends and achievement based on the 1986 national assessment* (Report No. 17-S-01). Princeton, NJ: Educational Testing Service.

Oliver, J.S., & Simpson, R.D. (1988). Influences of attitude toward science, achievement motivation, and science self concept on achievement in science: A longitudinal study. *Science Education, 72*, 143-155.

Parker, L.H., & Rennie, L.J. (1986). Sex stereotyped attitudes about science: Can they be changed? *European Journal of Science Education, 8*, 173-183.

Ridley, D.R., & Novak, J.D. (1983). Sex-related differences in high school science and mathematics enrolments: Do they give males a critical headstart toward science- and math-related careers? *Alberta Journal of Education, 29*, 308-318.

Sadker, D., & Sadker, M. (1986). Sexism in the classroom from grade school to graduate school. *Phi Delta Kappan, 56,* 512-515.

Sadker, D., Sadker, M., & Klein, S.S. (1986). Abolishing misconceptions about sex equity in education. *Theory into Practice, 25*, 219-226.

Science Council of Canada. (1984). *Science for every student: Educating Canadians for tomorrow's world*. Ottawa: Author.

Scott, J. (1982). Is there a problem? In J. Ferguson (Ed.), *Who turns the wheel?* (pp. 21-44). Ottawa: Science Council of Canada.

Simpson, R.D., & Oliver, J.S. (1990). A summary of major influences on attitude toward and achievement in science among adolescent students. *Science Education, 74*, 1-18.

Statistics Canada. (1988). *Canada yearbook: 1988*. Ottawa: Supply and Services Canada.

Steinkamp, M.W., & Maehr, M.L. (1984). Gender differences in educational motivations toward achievement in school science: A quantitative synthesis. *American Educational Research Journal, 21*, 39-59.

Taylor, H., Hunt, R., Sheppy, J., & Stronck, D. (1982). *British Columbia science assessment: General report*. Victoria, BC: Ministry of Education.

Tobin, K., & Gallagher, J.J. (1987). The role of target students in the science classroom. *Journal of Research in Science Teaching, 24*, 61-75.

Tobin, K., & Garnett, P. (1987). Gender related differences in science activities. *Science Education, 71*, 91-103.

Walberg, H.J. (1969). Physics, femininity, and creativity. *Developmental Psychology, 1*, 47-54.

Whyte, J. (1984). Observing sex stereotypes and interactions in the school lab and workshop. *Educational Review, 36*, 75-86.

Wideen, M.F. (1974). Division III science: An in depth evaluative study. *Saskatchewan Journal of Educational Research and Development, 5*(1), 15-20.

Wienekamp, H., Jansen, W., Fickenfrerichs, H., & Peper, P. (1987). Does unconscious behaviour of teachers cause chemistry lessons to be unpopular with girls? *International Journal of Science Education, 9*, 281-286.

Yager, R.E., & Yager, S.O. (1985). Changes in perceptions of science for third, seventh, and eleventh grade students. *Journal of Research in Science Teaching, 22*, 347-358.

Zarega, M.R., Haertel, G.D., Tsai, S., & Walberg, H.J. (1986). Late adolescent sex differences in science learning. *Science Education, 70*, 447-460.

Zimmerer, L.K., & Bennett, S.M. (1987, April). *Gender differences on the California statewide assessment of attitudes and achievement in science*. Paper presented to the Annual Meeting of the American Educational Research Association, Washington, DC.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 3, September 1991, 209-224*

JEFFREY W. BULCOCK
Memorial University of Newfoundland

MORLEY E. WHITT
E. Vaders Pentecostal College, St. John's

and

MONA J. BEEBE
Memorial University of Newfoundland

Gender Differences, Student Well-Being and High School Achievement

The study focuses on the question: Why are female high school students in Newfoundland outperforming boys in all school subjects except physics? Support for two theories—a biologically based theory and a culturally based theory—is reviewed and the arguments in favor of both are rejected. Instead, a demographic or cohort size hypothesis was entertained that generated three hypotheses: namely, that under circumstances when relative incomes are low (a) gender effects on the quality of school life will favor girls, (b) gender effects on student well-being will be mediated by the quality of school life and will favor girls, and (c) gender effects on school achievement will be mediated by the quality of school life and will, again, favor girls. While support for the demographic model of gender differences in school achievement was equivocal, the results indicated that student perceptions of their school lives influenced both their well-being and their achievement in literacy and numeracy.

Introduction

This study addressed the question: Why are Newfoundland girls outperforming Newfoundland boys in virtually all school subjects? For the past four years these gender differences have occurred at the grade 4, 6, 8, and 12 levels in the Newfoundland school system (Galway & Banfield, 1991, pp. 22-24; Department of Education, 1988, 1989, 1990, 1991).

For nearly a century there has been continuing research into the nature and origins of gender differences in achievement and ability. As far back as 1873, Herbert Spencer, in an article entitled "Psychology of the Sexes," argued in support of the theories of Charles Darwin that the intellectual attributes of women developed differently in the course of evolution (cited in Keeves, 1985). According to Sadker, Sadker, and Klein (1991), this turn of

Jeffrey Bulcock is a professor of education specializing in the social demography of education and research methodology.

Morley Whitt is a teacher in mathematics and physics with a research interest in issues in educational administration.

Mona Beebe is an associate professor of reading and language arts education. Her research interests include reading comprehension in the early grades, adult literacy, and the evaluation of reading instruction methodologies.

the century view that biology determined ability and personality implied that the range of roles open to women was more limited than that open to men (p. 307). In other words, although personality consists of the internalization of roles, roles themselves were subject to organic constraints that happened to favor men. In 1906, E.L. Thorndike rejected the view that gender differences in ability could be inherent on the grounds that the observed differences were too small to be of practical significance (cited in Keeves, 1985). Hollingworth, a student and colleague of Thorndike's, contended that the small differences observed were due to social influences and not to biological causes, and that the true intellectual potential of women would only be revealed when women received a similar education and had the right to choose equivalent careers (Hollingworth & Montague, 1914-1915). These views were congruent with those of the statistician Pearson (1897) who found few significant differences between men and women on a range of variables.

During the 1920s and 1930s research into gender differences was primarily concerned with the study of differences in attitudes and emotional needs. An important shift in research emphasis began in the 1950s. The assumption that roles were biologically given gave way to the thesis that the concept of role was integrally related to the patterns of behavior associated with a given status and position. Status and position in turn consist of expectations or systems of rights and duties (Sarbin, 1954, pp. 224-225). Thus to change roles was to change social expectations or the constellation of rights and duties associated with societal positions.

Maccoby's (1966) synthesis of research on gender differences in the United States, along with those of Tyler (1956) and Anastasi (1958), found that boys generally showed stronger numerical and spatial abilities and performed better on tests of mathematical reasoning than girls, but girls usually did better in verbal and linguistic studies. Later, Tyler (1969) specifically pointed out that when batteries of achievement tests were used to assess achievement, girls exceeded boys in performance in language studies and boys tended to perform better in mathematics and science. She emphasized, however, that these differences were not large, especially when compared with the differences within each gender group.

These findings from the 1960s were reinforced by Maccoby and Jacklin (1974) a decade later. It appeared, paradoxically, that while gender differences in the conative and affective domains were myths, gender differences in the cognitive domain were not. Elaborate explanations were formulated and tested to account for differences in intellectual performances. Especially prominent were hormonal theories (Geary, 1989; Resnick, Berenbaum, Gottersman, & Bouchard, 1986) and theories of hemispherical dominance (Dimond & Beaumont, 1974; Seward & Seward, 1980; Springer & Deutsch, 1981; Wheatley, Mitchell, Frankland, & Kraft, 1978).

The biological basis of sex role differences in cognition still remains a viable explanation. Recent meta-analytic reviews of the gender differences literature in mathematics by Hyde (1981), Feingold (1988), and Marsh (1989), however, show that the differences are small and declining. Ethington (1990), in an analysis of grades 7 and 8 mathematics data from eight countries, concluded that there were no gender effects and, further,

that what effects were discernible were in favor of girls. The same trend seems to hold in the case of science. In their meta-analytic review of the literature on gender differences in science, Linn and Hyde (1989) concluded that the gender gap is disappearing. They argue that gender differences in science reflect parallel differences in course enrollment and training, and that these differences are consistent with changing opportunity, changing social roles, and changing demands in the work place. Hollingworth's turn of the century prediction is being verified. In 1981, gender differences in math and science were the focus of a workshop entitled "Who Turns the Wheel?" dealing with the science education of women in Canada. The report on the workshop (Ferguson, 1982) indicates that studies in other countries have suggested that the separation of girls from science and mathematics starts early in life and that the gap becomes more obvious at each succeeding stage of the education process. As a result, few women graduate from university science and engineering faculties or from polytechnical institutes because they are not entering them. It may be that women do not enter these faculties because they are receiving inadequate science or mathematics training in high school. And this in turn may be a consequence of inadequate preparation and encouragement in the science areas at the primary and elementary school levels.

Physics courses in high school seem to have the least number of girls enrolled of any science/math course. Chemistry is studied more often than physics, and in biology courses there are often more girls than boys. Biology tends to be perceived by students as the least difficult science, and so it is taken by many girls merely to satisfy the science requirement for graduation. In advanced mathematics and pure mathematics courses often less than one third of the class are girls. In addition, by the final year of high school as many as one quarter of the female population in some provinces are no longer studying mathematics and less than half study physics or chemistry at that level. In contrast, 96% of the boys take mathematics in their final year of high school (Ferguson, 1982). Nevertheless, Bulcock and Mendoza (1988) found that when girls enroll in the more difficult science and math courses, they do as well as and sometimes better than boys. The interpretation of these data is complicated by the fact that the small percentage of girls who enroll in physics and chemistry may also represent the most academically able group. The larger percentage of boys is likely to include a significant proportion of individuals of more than average ability. Nonetheless, the achievement figures suggest that girls are capable of doing well in physics, chemistry, biology, and mathematics (Linn & Hyde, 1989).

Most of the data available on gender differences in achievement for Newfoundland are reported in the Department of Education statistical reports for the years 1988 through 1991. At grades 4 (Department of Education, 1991, Table 26), 6 (Department of Education, 1989, Figure 10), and 8 (Department of Education, 1990, Figure 14), girls outperformed boys in virtually all skill areas. These include reading comprehension, mathematics, and overall composite scores. The exception seems to be vocabulary at grade 8.

These findings prevail in grade 12 subjects with the exception of physics (Department of Education, 1988, Table 20; 1989, Table 23; 1990, Table 36). Although many of these differences are modest and therefore unlikely to be statistically significant, the differences in the verbal subjects such as language, literature, the New Testament, and French are substantial. It is necessary to point out, however, that the differences may be underestimated because there are, again with the exception of physics, more female students writing public examinations in these subjects than male students. For example, in the case of Language 3101, a compulsory subject, 4,121 female students wrote the final examination in 1989 compared with 3,517 male students (Department of Education, 1990, Table 36). If the top 3,517 female students were compared with the same number of male students, average achievement score differences would undoubtedly be higher than those found in the Department of Education report.

This section of the article presents two theories that purport to explain gender differences in school achievement. The first theory holds that cognitive behaviors are constrained by innate psychological or organic dispositions, whereas the second accounts for gender differences on the basis of cultural and social influences. However, the Newfoundland data on school achievement are at odds with traditional findings regarding gender differences and, because girls outperform boys in just about every subject in grades 4, 6, 8, and 12, the data do not support the traditional neo-Darwinist explanations. Further, there appears to be no obvious cultural or sociological explanation for these gender differences.

In the remainder of this article, the gender gap reversal in school achievement summarized above is reexamined from a demographic perspective followed by the testing of selected hypotheses derived from this perspective. It is emphasized from the outset that the use of a demographic perspective is exploratory and that the models formulated and tested are conjectural rather than confirmatory. Other influences such as changing societal expectations, expanding educational opportunities for women, and the perceived need for two incomes in families are certain to have contributed to the reversed gender differences observed in Newfoundland.

Demographic Perspective

The demography of human motivation, life style, income levels, and other social variables was developed by Easterlin (1980, 1987a, 1987b) over the past three decades and is referred to as the cohort size hypothesis. This theory holds that the fortunes of a birth cohort (those born in a given year) tend to vary inversely with its relative size; that is, if the birth rate is high adverse social effects will follow as a consequence of demographic crowding. The three institutions most affected by crowding are the family, the school, and the labor market.

During the depression years of the 1930s, birth rates were low. When this generation entered the labor market in the period 1950-1960, the supply of younger workers relative to older workers was low. This ratio of younger workers to older workers governs relative income such that when the ratio is low, relative incomes will be high. When the relative incomes of young men

(compared to the past incomes of their fathers) rise, they feel freer to marry and have children early in adulthood. If relative incomes fall the attendant economic distress will be reflected in declining marriage and fertility rates.

The evidence shows that the impact of declining relative income will be greater for those with high levels of education—for example, those with college degrees—hence there is less incentive for young men to invest in intellectual competencies when relative incomes are low. In contrast, when relative incomes are high, investments in schooling are profitable for young men. But high relative incomes promote high marriage rates and, because young men marry younger women, some women may forgo college for marriage, thereby retarding the growth of female college enrollments to some degree. The dynamics of these relationships between cohort size effects, relative income, and marriage rates account for the paradoxical situation that when incentives for investments in intellectual competencies for men are low they are high for women. In other words, when marriage is delayed, female college enrollments tend to accelerate because women have higher incentives for investing in schooling.

These relationships played themselves out in Canada in the post-World War II period. The baby boom peaked in Canada in 1959 with a total fertility rate of 3.9 (Statistics Canada, 1984). When those born in the 1950-1960 period entered the labor market in the 1970s, the supply of younger workers exceeded demand, and relative incomes declined throughout the subsequent decade of the 1980s. During this period, for the first time in Canadian history, the number of young women entering and graduating from college with first degrees exceeded the number of young men (Statistics Canada, 1991, Table 2.3.1.2). In the United States, which has had a similar post-World War II demography to that of Canada, the 1980s also marked a period during which female college enrollments exceeded male college enrollments (National Center for Education Statistics, 1989, Table 148).

Today the Canadian demographic scene is one of decline. Total fertility rates are at 1.7 (Press, 1990, Table 3), well below the 2.1 replacement level. Newfoundland is somewhat anomalous, however, in that when Canada had its baby boom Newfoundland had what can only be described as a super boom. The total fertility ratio in Canada peaked in 1959 at 3.9, but in Newfoundland it peaked in 1961 with a rate of 5.3 (Press, 1986, Figure 1). Today, Canada's total fertility rate is at an all-time low, 1.7; in Newfoundland it is even lower, 1.5 (Press, 1990). In fact, Newfoundland's total fertility rate has dropped from the highest of any province in Canada to the lowest in just 12 years. Today, the population of Newfoundland is in absolute decline.

As a result of the high fertility in Newfoundland between 1957 and 1963, these babies, now adults, have flooded the labor market, resulting in low relative incomes and high unemployment rates (in excess of 20%). Hence there are low marriage rates and declining incentives for young men to invest in schooling. In contrast, the low marriage rates provide opportunities for young women to complete high school and to attend postsecondary institutions, that is, to increase their educational investments. One would predict, then, that the declining incentives for young men to invest in schooling

concomitantly with the increasing opportunities for women to invest might be expressed as differential achievements—higher achievements for girls than for boys.

The Model

In the light of the above discussion and observations, it was hypothesized that the gender differences in school achievement in Newfoundland, which are at odds with those in other educational jurisdictions in both Canada and the United States, could be explained in terms of unusually volatile demographic shifts. If, as is argued above, changes in relative incomes affect incentives to invest in intellectual competencies, it might reasonably follow that when conditions favor investments for girls, but not for boys, perceptions of the quality of school life would prove more positive for girls than boys. Similarly, one might conclude that perceptions of the quality of school life would govern both satisfaction and dissatisfaction with schooling. In this way quality of school life would mediate the effects of gender on student well-being. There is also the outstanding question of whether the perceived quality of school life will also mediate the gender-school achievement relationships.

Thus far, three basic hypotheses have been derived from the above interpretation of Easterlin's hypothesis, namely, that under societal conditions where relative incomes are low: (a) gender effects on the perceived quality of school life will favor girls; (b) gender effects on school achievement will be mediated by the quality of school life; and, (c) gender effects on overall student well-being will be mediated by the quality of school life. The question arises whether the hypotheses account for both cognitive and affective outcome differences at all grade levels. Since the developmental acceleration of girls throughout childhood seems to apply to both intellectual as well as physical development (Anastasi, 1958), girls tend to be in advance of boys at every age from kindergarten through puberty or early adolescence. Since this presumed developmental advantage ends at puberty, the maturity hypothesis would no longer be applicable from approximately grade 8 through the remainder of the high school grades. For this reason the sample chosen for this study was a grade 10 sample of mostly 15-16-year-olds.

But by this grade level the average age of boys is higher than the average age for girls due largely to three factors: early school leaving, grade repeating, and the special education selection process. The data for the 1988-1989 school year show that 24% of the students who were in grade 8 four years earlier (1984-1985) had withdrawn from school early and that of these 63% were boys (Department of Education, 1989, Table 6). In the 1989-1990 school year 1,779 students in grades 1 through 9 repeated a grade (Department of Education, 1989, Table 17). This was 2% of the student body and since grade repeating occurs every year, as much as 18% of the school population repeats at least one grade. Again, about 63% of grade repeaters are boys (Department of Education, 1989, Table 18). Similar findings exist with respect to special education placement. May, Boone and Hopkins (1988) report that 66% of the special education students in the province are boys. It is largely for these reasons that boys are underrepresented in the senior high school grades (10-12) and are on average older than the girls. Consequently,

the analytical models called for by the three hypotheses include age as a covariate or exogenous variable. Two additional control variables were also considered—socioeconomic status and urban-rural location of the school attended by the sample respondents. Socioeconomic status was included because the resources of families are powerful determinants of educational behaviors (Hauser, 1971). Urban-rural location was included because rural students in Newfoundland consistently fall behind their urban peers in terms of school achievement (Department of Education, 1991, Table 28 and Figure 12). Thus, following the Simon-Blalock approach to causal modeling (Blalock, 1962; Simon, 1952), the variables socioeconomic status and urban-rural location were added to the exogenous variable set. The full model that was tested is illustrated in Figure 1.

Instrumentation

Data were obtained from a test battery, a student questionnaire, a parent questionnaire, and school records. The reading comprehension and mathematics scores at the end of the grade 10 year were obtained from teacher administered subtests of the Canadian Tests of Basic Skills, High School Battery (CTBS, King-Shaw, 1989). The quality of school life instrument was a version of the instrument developed by Williams and Batten (1981) and which was modified by Roberts and Clifton (1991). This instrument includes 29 items designed to measure four dimensions of the quality of school life—opportunity to learn in school (8 items), adventure in learning or intrinsic interest in schooling (7 items), perception of status as a student (7 items), and students' perception of their teachers (7 items). In addition, the quality of school life instrument contained seven schooling satisfaction items

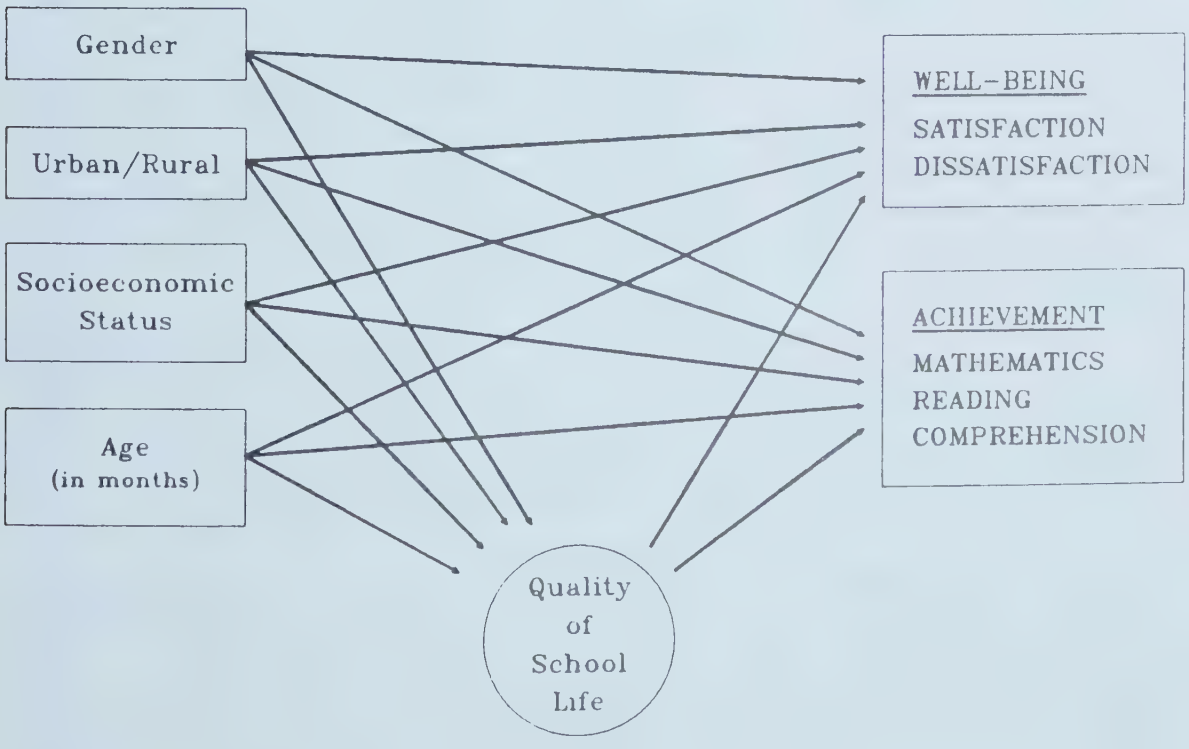


Figure 1. Conceptual model.

measuring positive affect and six schooling dissatisfaction items measuring negative affect.

Information on whether a school was located in an urban or rural community was obtained from Department of Education records. Communities with fewer than 5,000 people in the 1986 semi-decennial census were classified as rural. Information on parental education and occupation were obtained from a questionnaire completed by parents. Finally, the age and gender of the students were obtained from school records.

Sample

The population of schools with grade 10 classrooms was stratified by community size (urban and rural) and by denomination (Integrated and Catholic). Schools were chosen randomly from each cell. Schools located within the two school boards that refused to participate were replaced by "reserve" schools from the corresponding category. The sampled schools enrolled 450 grade 10 students (4.4% of the Newfoundland grade 10 population). Of these students, 317 (70.4%) received the informed consent of their parents to participate. Of those 317,182 (57.4%) were girls.

Measurement

Principal component analyses were conducted in order to construct weighted composites for seven variables—socioeconomic status, opportunity, adventure, status, teacher, satisfaction, and dissatisfaction. Because it was hypothesized (Williams & Batten, 1981) that opportunity, adventure, status, and teacher were dimensions of the quality of school life, they were used as indicators in a second-order factor analysis. It was also hypothesized that satisfaction and dissatisfaction were the dimensions of the student well-being construct. However, even though a well-being measurement model was estimated, it was not used except to estimate concurrent validity. It is argued here that when first-order composites are correlated with a second-order factor composite, the correlation is in fact an estimate of concurrent validity. In this case the concurrent validity is a measure of the relationship between first-order composites as indicators and the second-order composite as the criterion variable. The alpha reliabilities and concurrent validities of the composite variables are presented in Table 1.

The internal consistency reliabilities of the reading and mathematics achievement variables were not included in Table 1. In unreported analyses of CTBS results for selected Newfoundland grade level populations, the internal consistencies have been found to be marginally higher than those reported in the CTBS manuals (.90 ±.02, King-Shaw, 1990).

Results

The research hypotheses are embedded in the conceptual model presented in Figure 1. Model parameters for the five equations were estimated using ordinary least-squares regression: (a) quality of school life (QSL) regressed on the four exogenous variables (gender, urban-rural school location, socioeconomic status, and age); (b) schooling satisfaction regressed on the exogenous variables plus quality of school life; (c) schooling dissatisfaction regressed on the exogenous variables plus quality of school life; (d) mathe-

matics achievement regressed on the exogenous variables plus quality of school life; and (e) reading comprehension regressed on the same five variables as in (b) through (d). Equation (a) results are presented in Table 2. The estimates for equations (b) and (c) are presented in Table 3 and the estimates for equations (d) and (e) are presented in Table 4.

Because the metric for most model variables has little common language meaning, unstandardized betas are not reported. The effect parameters are standardized partial regression coefficients or path coefficients.

Discussion

In extending the cohort size hypothesis we argue that because prevailing demographic conditions in Newfoundland favor low relative incomes for young adults, the motivation to invest in intellectual competencies will be higher for female students than for male students. All the contemporary statistical evidence from official government sources supports the gender differences argument. High school girls in Newfoundland are performing as well as or better than high school boys in all school subjects. Similar trends for mathematics and science achievements have been reported in the United States.

It would seem to follow, therefore, that because girls are benefiting more than boys from their schooling opportunities, they would perceive both the quality of their school lives and their overall well-being as students in more

Table 1
Some Psychometric Properties of the Measurement Models

Measurement Model	Mean	Standard Deviation	N Items	Alpha Reliability	Concurrent Validity
SES (Socioeconomic Status)	49.73	10.00	2	.72	–
OPP (Opportunity to Learn)	50.00	10.00	8	.82	.83
ADV (Adventure in Learning)	50.00	10.00	7	.75	.77
ST (Student Status)	50.00	10.00	7	.71	.73
TCH (Perception of Teachers)	50.00	10.00	7	.80	.80
QSL (Quality of School Life)	50.00	9.45	4	.80	.67
SAT (Schooling Satisfaction)	50.00	9.21	7	.86	*
DIS (Schooling Dissatisfaction)	50.01	9.55	6	.70	–.562**

*Schooling satisfaction was used as a criterion for the QSL composite.

**Represents the convergent validity between schooling satisfaction and schooling dissatisfaction.

Table 2
Correlations, Direct Effects and t-Values for the Effects of Exogenous Variables on Quality of School Life

Intervening Variable	Exogenous Variables	Correlation (r)	Direct Effect	t-Value[a]
QSL				
	AGE	.035	.065	1.141
	GENDER	.092	.110	1.931
	UR	−.024	−.002	−.041
	SES	.086	.107	1.838

Multiple R = .14715
R^2 = .02165
[a]The t-value is for the direct effects only.

Key: AGE = Age in Months
GENDER = Male (1), Female (2)
UR = Urban (2), Rural (1)
SES = Socioeconomic Status
QSL = Quality of School Life

favorable terms than boys. Models designed to validate these propositions were formulated and estimated. The results are presented in Tables 2, 3, and 4. The findings were equivocal.

First, gender differences in the perceived quality of student life reported in Table 2 were not significant when controlling for age, urban-rural residence, and socioeconomic status differences. Second, gender difference in terms of their direct effects on schooling satisfaction and schooling dissatisfaction were not significant either (Table 3). When the effects of the gender model of schooling satisfaction were decomposed following Alwin and Hauser (1975), we were able to show that the total effect of gender on schooling satisfaction (Table 3) as mediated by the perceived quality of school life was substantially higher than the direct effect (0.146 compared with .073). The question arises, then, as to whether the ratio of the estimator to its standard error follows the *t* distribution (Folmer, 1981; Sobel, 1982; and Bollen, 1987). If it does, then the statistical significance of total effects can be estimated by the classical method. The jury is still out on this question, although recent findings (Bollen & Stine, 1990) show that in moderately large samples the distribution of an estimator is close to that assumed by the classical method. And since the sample size in the present study is large (N=317) we are able to assume that a total effect parameter of 0.146 will generate (using the classical method) a t=2.619 ($p < .05$). The same cannot be concluded about the total effect of gender on schooling dissatisfaction (TE=−.028). We conclude, therefore, that in terms of overall schooling satisfaction girls are generally more satisfied than boys, whereas in terms of overall schooling dissatisfaction there are no discernible gender differences. There is support for the proposition that girls find their schooling experiences more satisfying than do boys. But there is no support for two other propositions, namely, that girls will report higher quality school lives and less schooling dissatisfaction

Table 3
Correlations, Direct Effects, Indirect Effects, Total Effects and t-Values for the Effects of Independent Variables on the Dimensions of Student Well-Being

Outcome Variables	Independent Variables	Correlation (r)	Direct Effect	Indirect Effect	Total Effect	t-Value[a]
Satisfaction						
	AGE	.010	−.004	.043	.039	−.096
	GENDER	.134	.073	.073	.146	1.713
	UR	−.022	−.005	−.002	−.007	−.117
	SES	−.057	.006	.071	.077	.131
	QSL	.669	.661	—	.661	15.595*
Multiple R = .67272 R^2 = .45255						
Dissatisfaction						
	AGE	−.139	−.117	−.036	−.153	−2.454*
	GENDER	−.001	.033	−.061	−.028	.701
	UR	.063	.042	.001	.043	.891
	SES	−.062	−.022	−.059	−.081	−.447
	QSL	−.556	−.552	—	−.552	−11.734*
Multiple R = .57183 R^2 = .32699						

[a]The t-value is for the direct effects only.

Key: AGE = Age in Months
GENDER = Male (1), Female (2)
UR = Urban (2), Rural (1)
SES = Socioeconomic Status
QSL = Quality of School Life

*$t \geq 2.0$ is significant at the $p \leq .05$ level.

than boys. That there was an absence of support for either of these latter propositions calls for an ex post facto explanation.

It is explained above that the sampling frame for this study was based on grade, not age. Age-based samples are troublesome for the school authorities to obtain because they cross several grade levels. To avoid refusals to participate, grade-based samples are the compromise solution. But grade is not a good proxy for age at the high school level. Early school withdrawal, grade repeating, and special education arrangements distort the age-gender distributions of grade 10 classrooms because about two thirds of the members of these categories are boys. Two important consequences follow. First, there will be more female than male students in grade 10 and, second, the average age of male students will be higher than the average age of female students. In the present sample the correlation between age and gender was −0.121 ($t=-2.213$, $p < .02$), showing that older students were more likely to be male than female. Because it is plausible to assume that the absent male students

Table 4
Correlations, Direct Effects, Indirect Effects, Total Effects and t-Values for the Effects of Independent Variables on Reading and Mathematics

Outcome Variables	Independent Variables	Correlation (r)	Direct Effect	Indirect Effect	Total Effect	t-Value[a]
Reading Comprehension						
	AGE	−.227	−.190	.013	−.177	−3.770*
	GENDER	.054	.040	.022	.062	.798
	UR	−.261	−.210	.002	−.210	−4.199*
	SES	.366	.285	.021	.306	5.550*
	QSL	.224	.198	—	.198	3.983*
Multiple	R = .49894					
	R^2 = .19452					
Mathematics						
	AGE	−.264	−.268	.011	−.257	−5.119*
	GENDER	−.139	−.170	.019	−.151	−3.281*
	UR	−.176	−.150	.000	−.150	−2.885*
	SES	.272	.172	.018	.190	3.228*
	QSL	.162	.169	—	.169	3.277*
Multiple	R = .44105					
	R^2 = .24894					

[a]The t-value is for the direct effects only.

*t $\geq$ 2.0 is significant at the $p \leq .05$ level.

were absent at least partially because they were ambivalent about schooling, the fact that they were absent would likely bias the findings in favor of support for the null hypothesis.

This explanation for the equivocal findings of the gender-schooling satisfaction relationship may also account for the lack of support for the thesis that gender effects on school achievement will favor girls. As shown by Table 4 results, there were no gender effects on reading comprehension, and the effects of gender on mathematics achievement was in favor of grade 10 boys. Again, these results are congruent with a sample bias in which girls were overrepresented.

Other features of the Figure 1 model are worth noting. In terms of the well-being equations reported in Table 3, none of the control variables was significant. Whether grade 10 students are satisfied or dissatisfied with their school careers depends largely on their perceptions of the quality of their school lives. Those reporting a low quality of school life tended to be the least satisfied and the most dissatisfied with schooling. Since these models had a good fit, with R-square values greater than 0.3, they seem worth taking

seriously. We know that there are four dimensions to the quality of school life construct and that school authorities might be able to manipulate these dimensions (opportunity to learn, adventure in learning, student status, and student identity) for the better, thereby bringing about enhanced student well-being. What is just as important, however, is that the quality of school life construct has significant positive effects on cognitive performance as measured by reading comprehension and mathematics ability. And these effects are over and above the effects of four control variables—age, gender, urban-rural residence, and socioeconomic status.

Note, however, that by grade 10 the debilitating effects of social background (SES) and the constraints of the rural environment on school achievement have already occurred—at least in Newfoundland. And note that, as predicted, age effects are both negative and statistically significant. The older students in grade 10 are the weakest students. In other words, by grade 10, age in the Newfoundland school system is to some extent a proxy for grade repeating and developmental delay.

Finally, we comment on the quality of student life model reported in Table 2. Our attempt to explain student perceptions of the quality of school life on the basis of selected demographic variables—age, gender, urban-rural residence, and socioeconomic status—was unsuccessful. These factors accounted for a mere 2% of the variance. Age, urban-rural residence, and socioeconomic status did not account for the quality of school life and, just as important, they did not account for schooling satisfaction either. Age, paradoxically, was a significant factor in accounting for schooling dissatisfaction (Table 3). Older students in grade 10, other things equal, were less dissatisfied with schooling than younger students, a finding for which we offer no explanation.

Conclusion

This article focuses on gender differences in student achievement in a province where girls are performing as well as and mostly better than boys in all school subjects at the high school level. Two theories purporting to explain gender differences in achievement are reviewed, namely, organic or endogenous psychological dispositions on the one hand, and cultural or social influences on the other. Neither seemed capable of accounting for the distinctive Newfoundland data. Instead, a cohort size hypothesis was examined that at the level of descriptive analysis seemed to provide a plausible explanation for gender differences in achievement in favor of girls. It generated the hypothesis that when relative incomes are low, male incentives to invest in intellectual competencies are low while, in contrast, when relative incomes are low, female incentives and opportunities are high. It followed from this that, if demographic circumstances were favorable for girls but not for boys, then the relationship between gender and both the perceived quality of student life and student well-being would favor girls.

These are dynamic relationships in that they are responsive over time to changing cohort size and relative incomes. Macro data suitable for the estimation of dynamic time series models were not available. Instead, static models based on cross-sectional data had to be used as the best substitute.

Not surprisingly, therefore, these exploratory micro-models with students as the units of analysis tended to be conjectural and the results somewhat equivocal. Nevertheless, total effect parameters in one model supported the thesis. That stronger support was not forthcoming may have been caused by unavoidable sampling design flaws that involved using grade level rather than age as the sampling frame. This limitation of the design generated a biased sample even though every effort was made to use appropriate statistical controls.

These limitations indicate the problems in testing the cohort size hypothesis as an explanation of gender differences in school achievement. The focus on gender differences, however, should not detract from some of the other interesting relationships found between student perceptions of the quality of their school lives and their overall student well-being. It was argued that the dimensions of the quality of school life construct were potentially policy manageable. These dimensions proved not only powerful predictors of well-being (schooling satisfaction and dissatisfaction), but also powerful predictors of literacy (reading) and numeracy (mathematics). School authorities might note, therefore, that increased emphasis on schooling quality could have important payoffs in terms of both schooling satisfaction and achievement for high school students.

Acknowledgment

This study was funded by the Department of the Secretary of State of Canada. The research assistants on the project were Shelley Hasinoff and Marguerite Baker. Jill Andrews and John Humber, Division of Evaluation and Research, Department of Education, Government of Newfoundland provided additional assistance.

References

Alwin, D.F., & Hauser, R.M. (1975). The decomposition of effects in path analysis. *American Sociological Review, 40*, 37-47.

Anastasi, A. (1958). *Differential psychology: Individual and group differences in behavior* (3rd ed.). New York: Macmillan.

Blalock, H.M. (1962). Four-variable causal models and partial correlations. *American Journal of Sociology, 68*, 182-194.

Bollen, K.A. (1987). Total, direct, and indirect effects in structural equation models. In C.C. Clogg (Ed.), *Sociological methodology* (pp. 37-69). Washington, DC: American Sociological Assosication.

Bollen, K.A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. In C.C. Clogg (Ed.), *Sociological methodology* (pp. 115-140). Oxford: Blackwell.

Bulcock J.W., & Mendoza, L. (1988). *The academic standing of the 1987 education graduates at Memorial University*. St. John's, NF: Institute for Educational Research and Development, Memorial University of Newfoundland.

Department of Education. (1988). *Education statistics*. St. John's, NF: Division of Evaluation and Research, Government of Newfoundland and Labrador.

Department of Education. (1989). *Education statistics*. St. John's, NF: Division of Evaluation and Research, Government of Newfoundland and Labrador.

Department of Education. (1990). *Education statistics: Elementary-secondary*. St. John's, NF: Division of Evaluation and Research, Government of Newfoundland and Labrador.

Department of Education. (1991). *Education statistics*. St. John's NF: Division of Evaluation and Research, Government of Newfoundland and Labrador.

Dimond, S.J., & Beaumont, J.E. (Eds.). (1974). *Hemisphere function in the human brain*. New York: Wiley.

Easterlin, R.A. (1980). *Birth and fortune*. New York: Basic Books.

Easterlin, R.A. (1987a). Easterlin hypothesis. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: A dictionary of economics* (vol. 2, pp. 1-4). London: Macmillan.

Easterlin, R.A. (1987b). *Birth and fortune: The impact of numbers on personal welfare* (2nd ed.). Chicago, IL: University of Chicago Press.

Ethington, C. (1990). Gender differences in mathematics: An international perspective. *Journal for Research in Mathematics, 21*, 74-80.

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43*, 95-103.

Ferguson, J. (1982). *Who turns the wheel?* Ottawa: Publications Office, Science Council of Canada.

Folmer, H. (1981). Measurement of the effects of regional policy instruments by means of linear structural equation models and panel data. *Environment and Planning, 13*, 1435-1438.

Galway, G., & Banfield H. (1991). *Profile '90: Educational indicators*. St. John's, NF: Department of Education, Government of Newfoundland and Labrador.

Geary, D. (1989). Comment: A model for representing gender differences in mathematical tasks. *Review of Educational Research, 59*, 185-213.

Hauser, R.M. (1971). *Socioeconomic background and educational performance*. Washington, DC: American Sociological Association.

Hollingworth, L.S., & Montague, H. (1914-1915). The comparative variability of the sexes at birth. *American Journal of Sociology, 20*, 335-370.

Hyde, J.S. (1981). How large are cognitive gender differences? A meta-analysis using w and d. *American Psychologist, 36*, 892-901.

Keeves, J.P. (1985). Sex differences in ability and achievement. In T. Husen & T.N. Postlewaite (Eds.), *International encyclopedia of education* (pp. 4438-4544). Oxford: Pergamon Press.

King-Shaw, E. (Ed.). (1989). *Canadian tests of basic skills: Teacher's guide*. Scarborough, ON: Nelson Canada.

King-Shaw, E. (Ed.). (1990). *Canadian tests of basic skills: Manual for administrators, supervisors and counsellors*. Scarborough, ON: Nelson Canada.

Linn, M.C., & Hyde, J.S. (1989). Gender, mathematics and science. *Educational Researcher, 18*, 17-19, 22-27.

Maccoby, E.E. (Ed.). (1966). *The development of sex differences*. Stanford, CA: Stanford University Press.

Maccoby, E.E., & Jacklin, C.N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Marsh, H. (1989). Sex differences in the development of verbal and mathematical constructs: The high school and beyond study. *American Educational Research Journal, 26*, 191-225.

May, S., Boone, W., & Hopkins, B. (1988). *Sex-related differences in mathematics: Participation, achievement, and attitudes in the Newfoundland and Labrador setting*. St. John's, NF: Division of Evaluation and Research, Government of Newfoundland and Labrador.

National Center for Education Statistics. (1989). *Digest of education statistics*. Washington, DC: Department of Education, Office of Educational Research & Improvement.

Pearson, K. (1897). *The chances of death and other studies in evolution* (vol. 1). London: Arnold.

Press, H.L. (1986). *Trends in education: Demographic projections*. St. John's, NF: Division of Evaluation and Research, Government of Newfoundland and Labrador.

Press, H.L. (1990). *Toward 2000: Trends report II, elementary-secondary projections*. St. John's, NF: Division of Evaluation and Research, Government of Newfoundland Labrador.

Resnick, S.M., Berenbaum, S.A., Gottersman, I.I., & Bouchard, T.J. (1986). Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology, 22*, 191-198.

Roberts, L.W., & Clifton, R.A. (1991). *Measuring the quality of university student life*. Winnipeg, MB: Department of Sociology, University of Manitoba.

Sadker, M., Sadker, L., & Klein, S. (1991). The issue of gender in elementary and secondary education. In G. Grant (Ed.), *Review of Research in Education* (vol. 17, pp. 269-334). Washington, DC: American Educational Research Association.

Sarbin, T.R. (1954). Role theory. In G. Lindzey (Ed.), *Handbook of social psychology* (vol.1, pp. 223-258). London: Addison-Wesley.

Seward, J.P., & Seward, G.H. (1980). *Sex differences: Mental and temperamental*. Lexington, MA: C.D. Heath.

Simon, H.A. (1952). On the definition of the causal relation. *Journal of Philosophy, 49*, 517-528.

Sobel, M.E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290-312). San Francisco, CA: Jossey Bass.

Springer, S.P., & Deutsch, G. (1981). *Left brain, right brain*. San Francisco, CA: Freeman.

Statistics Canada. (1984). *Fertility in Canada: From baby boom to baby bust*. Ottawa: Minister of Supply & Services, Government of Canada.

Statistics Canada (1991). *A statistical portrait of university education in Canada*. Toronto, ON: Council of Ministers of Education Canada, Government of Canada.

Tyler, L.E. (1956). The psychology of human differences (2nd ed.). New York: Appleton-Century-Crofts.

Tyler, L.E. (1969). Sex differences. In R.L. Ebel (Ed.), *Encyclopedia of educational research* (4th ed., pp. 1217-1221). Toronto, ON: Macmillan.

Wheatley, G.H., Mitchell, R., Frankland, R.L., & Kraft, R. (1978). Hemisphere specialization and cognitive development: Implications for mathematics education. *Journal for Research in Mathematics Education, 9*, 20-32.

Williams, T., & Batten, M. (1981). *The quality of school life*. Hawthorn, Victoria: Australian Council for Educational Research.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 3, September 1991, 225-239*

GAALEN ERICKSON
University of British Columbia

and

SANDRA FARKAS
Vancouver Technical Institute

Prior Experience and Gender Differences in Science Achievement

A number of conjectures voiced in the literature concerning the role of prior experiences as a possible factor in accounting for the observed gender differences in science achievement tests were examined. A set of 23 multiple-choice items that displayed significant gender differences on a previous British Columbia science assessment test were administered to 238 grade 12 students. Fifteen of these students were selected for a follow-up interview to probe the reasons they gave for their responses to these items. In particular, they were encouraged to identify any previous experiences they might have had that might have enabled them to respond to the item. Four categories were generated from their explanations of why they responded as they did to the items. These were school-based experiences, informal experiences, common knowledge, and a final miscellaneous category. The results indicate that females consistently appealed to school-based experiences much more than did the males (even though this did not translate into better results on the items). A second finding was that many of the items elicited spontaneous negative reactions for many of the females and virtually none of the males. In examining these results and drawing on the rapidly expanding literature in this field of inquiry, it was concluded that efforts to remove the gender bias in achievement testing will require a systematic, empirical analysis of the experiential contexts of items.

Introduction

That the issue of gender effects in science achievement and enrollment patterns remains one of current interest can be readily inferred from the continued expansion of the research literature in this area (Bearlin, Annice, & Elvin, 1988; Kelly, 1987a; Koshland, 1988; Linn & Hyde, 1989). An even more important indicator of the vitality of the field in science education is the increasing institutionalization of resources aimed at developing effective policies and practices to address some of the problems of females and science that have been identified in the literature. National and international organizations have been created to sponsor conferences, workshops, and projects designed to create and disseminate materials and strategies for the

Gaalen Erickson is an associate professor in the Department of Mathematics and Science Education, and Associate Director of the Centre for the Study of Teacher Education. His research interests are in the areas of science education, gender issues in the assessment context, and teacher education.

Sandra Farkas is a teacher in the Vancouver School District. Her research interests are in gender equity and classroom practice.

teaching of science/technology subjects in a way that will not disadvantage females. Of the international organizations, the Girls and Science and Technology (GASAT) group has been the most active, having sponsored five international meetings in the past eight years.

On a smaller scale many action groups at provincial/state or even at local levels (universities or school districts) have been established for the purposes of developing gender-conscious curricula, organizing workshops and inservice programs for educators, and lobbying political organizations. Examples of many of these action groups all over the world have been reported at recent GASAT meetings (Daniels & Kahle, 1987; Harding, 1986; Raat, 1981; Sjoberg, 1983).

Although the proliferation of these groups does signal a recognition of the importance of gender issues, it does not mean that all of the questions concerning gender effects have been resolved. One area of continuing discussion in both the research literature and in the various programs is that of the nature of the possible sources of the observed gender differences. While much of the earlier debate in the 1970s focused on the roles played by biological versus sociological factors (Kelly, 1981; Kimball, 1982; Maccoby & Jacklin, 1974; Sherman, 1978), most of the recent research and program development has been directed toward identifying those sociocultural factors that contribute to the observed differences, particularly those that might be amenable to change by altering current school practices (Bentley & Watts, 1986; Crossman, 1987; Haggerty, 1987; Linn & Hyde, 1989; Martin, 1989; McClintock Collective, 1988).

One of the arguments that has been consistently made to account for girls' lower achievement scores in the physical science subjects is that in their growing years girls tend to have fewer and different kinds of experiences with a wide range of physical phenomena relevant to these subjects (Johnson & Murphy, 1984; Kahle & Lakes, 1983; Lie & Bryhni, 1983; Sjoberg, 1989). In some instances these arguments are advanced as intervening premises suggesting that girls' prior experiences could account for observed gender differences in spatial and mechanical reasoning scores that may in turn be the basis for differences in achievement scores (Finn, Dulberg, & Reis, 1979; Maccoby & Jacklin, 1974). In other cases the argument is constructed to implicate prior experience directly in accounting for the differences in standardized achievement tests or enrollment patterns (Erickson & Erickson, 1984; Johnson & Murphy, 1984; Kelly, 1987b; Parsons-Chatman, 1987). For example, Erickson and Erickson (1984) examined a number of items in a large-scale achievement test that displayed substantial gender differences and proposed that some of the observed differences could be explained on the basis of the item context. Those items that were embedded in a context that was much more familiar to males (i.e., were part of their "sphere of experience") may have conveyed an advantage to the males by allowing them to construct more plausible responses to the items even if they did not "know" the answer from their direct school experiences in science. For example, items that used balls, electrical circuits, compasses, and other materials or situations that are more likely to be a part of a boy's everyday experiences

seemed to have particularly large differences in achievement scores between males and females.

Bateson and Parsons-chatman (1989) went one step further in checking out the validity of this conjecture. They developed a scale, which they labeled *gender orientation,* to characterize items in terms of the degree to which they were associated with or related to the prior experiences of grade 7 girls or boys. They then asked 20 adults (including both elementary school teachers and parents) to use this scale to rate 120 items that were used in the 1986 British Columbia Science Assessment (Bateson, Anderson, McConnell, Rutherford, & Dale, 1986). Correlating the mean gender orientation rating for each item with the differences in the achievement scores between boys and girls on these same items, they were able to obtain an estimate of the relationship between these two variables. The correlation coefficient they reported for the entire 120-item test was .53 while the correlations for some of the subtests ranged from a low of .18 to a high of .84 (Bateson & Parsons-chatman, 1989). These data suggest that prior experience may indeed be important in influencing performance on achievement tests, but because Bateson and Parsons-chatman had to use judgments from adults to assign gender orientation scores, their empirical support for the prior experience argument is indirect. More direct investigations of the prior experience hypothesis, both to assess its validity and to begin sketching out its dimensions in qualitative terms, are clearly desirable.

Overview of the Study

The work reported here represents an initial endeavor to begin mapping out in more detail some of the relevant qualitative factors in this problem area. The study is clearly exploratory in nature and was designed to examine in a more interpersonal context the apparent impact of prior experience on students' performance on a multiple-choice science achievement test. The specific intent of the present study was to obtain more information about prior experience from the students themselves rather than using more indirect methods. The value of obtaining data directly from the participants regarding the meaning they attach to particular situations (such as responding to a multiple-choice item) has been underscored in the recent debates on competing traditions of educational inquiry (Erickson, 1986; Firestone, 1987; Howe, 1988; Marton, 1988; Smith & Heshusius, 1986). Since this study is aimed at uncovering some of the meaning that students attach to responding to science achievement items, it is located in the growing tradition of *interpretive research* (Soltis, 1984).

The Research Questions

Several specific questions were used to guide the collection and the analysis of the data for this study. These specific questions were:

1. What are the different kinds of explanations students give in accounting for their choices on a set of science achievement items? Can these explanations be categorized according to the type of prior experiences that they draw on?
2. Are there any gender differences in these explanatory categories?

3. Does the item context account for any gender differences?

Methods

Data Collection

Two related data bases were generated in the study. One was obtained from a sample of 238 grade 12 students selected from three schools in the Vancouver Public School system. These students were administered a 23-item, multiple-choice test (see Farkas, 1986, for a complete description of the items and the procedures employed). Table 1 provides a classification of the items used by goal area and content area. As shown, the items were distributed among the cells formed by crossing the goal and content areas. The 23 items had been used earlier in a larger, province-wide science assessment in 1982 (Taylor, Hunt, Sheppey, & Stronck, 1982) and had yielded differences in scores in favor of males ranging from 8 to 30 percentage points.

The items were administered to 238 students for three reasons. First, we wanted to ensure that there was some stability in the gender differences that had been observed in these items in the previous (1982) assessment. Second, we wanted to simulate as closely as possible the types of conditions (e.g., large group, limited time, and not in a science class) under which the BC Assessment measures are normally administered. Finally, we selected from this sample of students a smaller group to interview. We thought it was important that the students we interviewed had originally responded to the items under assessment-like conditions.

The test was administered by the regular classroom teachers to the students in their English or consumer education classes. (These were the only two compulsory grade 12 subjects at the time of the study.) Using these students enabled us to obtain a much more heterogeneous and representative sample than if we had used grade 12 science classes.

Table 1
Classification of Test Item Numbers by Goal and Content Area

	Content Area			
Goal Area	Biology	Chemistry	Physics	Earth/Space
I. Science Processes				
a. interpret data	20,21[a]		1,2	
b. identify and control variables		11		
II. Knowledge, Recall				
a. understanding	9	18	6	5,16
b. safety procedures		7,8,17,22	13	
III. Higher Level Thinking				
a. evaluate evidence for conclusions			15,23	12
b. solve abstract problems		10	14,19	3,4

[a]Item numbers

For our purposes the most important data for the study were those obtained from a series of semistructured personal interviews with 15 students (8 males and 7 females) selected from the larger group. These students were selected by their respective teachers on the grounds that they represented a typical cross-section of students in terms of academic performance and other background variables. The interviews, which were conducted by the second author within a week after the students had completed the original test, were designed to identify linkages between the students' self-reported prior experiences and their written responses to the multiple-choice items included in the test.

Each interview began with a brief orientation period developed to put the students at ease and to explain that we were interested in their ideas about how they arrived at their answers to the items and not whether they had the "right" answer. The students were then given their own unmarked test so that they could review each question and their response to it. The students were asked to reread each question and their response to it. They were then asked some variation of the following question: "Could you remember what you were thinking when you answered this question or can you explain why you selected this answer?" The interviewer often asked for further elaboration of their response depending on the nature of their initial response. These interviews, which typically lasted for 30 to 40 minutes, were recorded on audiotape for later transcription and analysis. Illustrative comments from some of these interviews are provided in the Results section.

Data Analysis

The data analysis of the student responses to the 23-item test consisted of scoring the items according to correct and incorrect answers. Analysis of the interview data focused on developing an appropriate set of categories to account for the explanations offered by the students in justifying their responses to a given item. First, a preliminary list of all the reasons offered by the students was prepared. These reasons, many of which were based on students' recollections of previous school experiences, were then grouped into four categories: school-based experiences, informal experiences, common knowledge, and miscellaneous. Each of these categories is described and illustrated in the presentation of results of the analysis of the students' responses.

Results

Written Tests

The percentage correct between the males and females and the difference in these percentages for both the 1982 assessment and the current administration are presented in Table 2 for each of the 23 multiple-choice items included on the test.

There were three items (8, 15, and 22) for which the difference scores varied by more than 10% between the two assessments. We have no easy explanation for these changes except to say that they were created from a sizable decrease in the performance by the males combined with a small increase in performance by the females. It is interesting to note in passing that items 8 and 22 assessed chemistry safety procedures. The issue of safety

Table 2
Percentage Scores for Selected Science Achievement Items From 1982 BC Science Assessment and Current Study by Gender

	1982 Results (N = 22,110)			Current Results (N = 238)		
Item No.	Male	Female	(M - F)	Male	Female	(M - F)
1	46%	38%	8%	49%	32%	17%
2	43	26	17	42	27	15
3	48	26	22	60	31	29
4	41	19	22	42	26	16
5	36	25	11	54	36	18
6	60	34	26	65	38	27
7	53	38	15	55	36	19
8	27	17	10	20	22	–2
9	71	57	14	72	48	24
10	50	32	18	57	32	25
11	57	46	11	68	52	16
12	52	38	14	52	45	7
13	74	44	30	66	35	31
14	76	55	21	72	51	21
15	69	44	25	54	47	7
16	58	39	19	57	39	18
17	80	67	13	67	51	16
18	72	62	10	73	61	12
19	69	46	23	70	50	20
20	47	37	10	41	32	9
21	66	58	8	58	40	18
22	70	46	24	60	51	9
23	60	45	15	63	51	12
Mean	57.6	40.8	16.8	57.3	40.6	16.7
SD	13.7	13.0	6.2	12.3	10.0	7.6

Note: Percentage scores are rounded off to the nearest whole number.

has recently received considerable attention in science classes, but this emphasis does not seem to be reflected in the males' performance on these two items. But in aggregate scores, the 1982 and current samples were virtually identical for both the item scores and the difference scores.

However, given the similarity of both the mean scores and the difference scores for both groups, we would argue that the gender differences in the responses to the items have remained relatively constant over the four-year interval between the two test administrations.

Interviews

The interview data suggest that the students appeared to be comfortable in discussing the reasoning and/or experiences that they used in selecting their responses. A wide range of comments and recounting of experiences was elicited from these interviews. An illustration of the type of comments made is provided in the following excerpt (I = Interviewer, S = 17-year-old female student):

> I: What were you thinking when you chose the answer you did for question 7? [A question on the proper procedure for lighting a Bunsen burner.]
>
> S: Okay. When I was in grade 9 I had a bad experience with a Bunsen burner. I don't know whether my answer is right or not. When I lit the Bunsen burner it kind of shocked me. My science teacher gave me a sheet and said that you should read it and study it and know where the air control should be. Well, I'm never going to light another Bunsen burner; I'm going to get someone else to do it so I never read it. So, what I put for my answer was number four because air allows a flame to burn. I didn't want to put "I don't know" because I kind of had a bit of knowledge about it.

Four explanatory categories were constructed from the student responses. The first, labeled school-based experiences, included any student comments that referred to some experience related to, or knowledge obtained from, some previous school activity. The excerpt given above, for example, was placed in this category. A second example is the comment: "Well, it's something I remember my teacher telling me to do—you always mix the water with the acid" that was provided by a female responding to an item asking for the appropriate laboratory procedure for diluting acid. It should be noted at the outset that as in the foregoing example an appeal to some school-based experience did not necessarily result in a "correct" answer.

The second category, informal experiences, was used to categorize student references to experiences or events such as hobbies, playing games, television, radio, and recreational reading. In other words, this category included experiences obtained outside the school context. For example, a male responded to a question on the purpose of circuit breakers by stating: "That I know from stuff around the house. Fuses blow in our house all the time. I usually go downstairs and switch the breakers."

Common knowledge was the third category, and it included references to knowledge that the student held where its source could not be readily identified. A number of students made statements such as "Well, that's just basic knowledge. Something I learned a long time ago" and "This sort of made sense to me. It was logical."

The final category, which we have simply labeled as miscellaneous, included such things as guessing at the answer, not being able to remember how they had obtained an answer, or working out the answer by a process of elimination. The types of comments included in this category are typified by the following female response: "Oh boy, I don't know what would happen. I have too much pride to put 'I don't know,' so I guessed." A slightly different

rationale was offered by this male student: "At the time I knew the answer, but it seems too long and complicated to remember what I was thinking."

In response to our first research question, then, we think these examples demonstrate that a semistructured interview does appear to be an effective means of identifying a wide range of prior experiences that the students claimed to have used in generating their responses to the written test. Furthermore, it was possible to develop a category system that enabled us not only to accommodate all of the students' responses, but, more importantly, provide some potentially fruitful conjectures about how males are able to draw on a wider repertoire of experiences in answering selected science questions.

The second question we examined was whether there were any gender differences in the responses provided by the students. One way of identifying these differences was to tabulate these explanatory responses by category and by gender. Table 3 provides these data. An inspection of this table reveals that the response category with the largest gender difference is that of school-based experiences. The females referred to such experiences for just over one half of their responses, while the males referred to school-based experiences for just over one third of their replies. In contrast, males more than females referred to other experiences as a basis for their responses. In brief, it seems clear that there were indeed consequential gender effects among the types of prior experiences offered by the students. The females more than the males relied more on information/experiences remembered from their school science, whereas the males appealed more than females to knowledge obtained from their out-of-school interests and what we have termed as common knowledge.

The final research question of this study was addressed in two ways. The first was to identify items where the context is most likely to be gender sensitive (Bateson & Parsons-chatman, 1989; Erickson & Erickson, 1984) and examine the students' explanatory responses for those items. The second approach was to identify those items that elicited a spontaneous negative reaction from the students and to determine the source of these reactions. It is noteworthy that all the students, with a single exception, who responded in this spontaneous manner to particular features in the question stems were females.

Table 3
Male and Female Explanatory Responses by Experience Category

	(n)	School-based Experiences	Informal Experiences	Common Sense	Misc.	Total No. of Responses
Male	8	66 (36)	38 (21)	40 (22)	40 (22)	184
Female	7	87 (54)	26 (16)	22 (14)	26 (16)	161
Difference (M - F)		(–18)	(5)	(8)	(6)	

Note: Percentage of explanations for all 23 items is given in brackets and rounded off to the nearest whole number.

A context that has been identified by most researchers in this field as being very gender sensitive is that of electricity (Johnson & Murphy, 1984; Kelly, 1987b; Parsons-Chatman, 1987). We found this was apparent in our study as well. Question 13, which generated the largest difference score (31%), asked about the purpose of a fuse or circuit breaker. In describing the reasons for their answers, six out of the seven females interviewed referred to some type of school-related experience to justify their selection. As indicated earlier, however, this attempt to remember relevant ideas or details from earlier school experiences often resulted in incorrect answers or confusion about the concept being assessed as illustrated by the following response:

> I know what a circuit breaker is. It makes things go in two directions. We did it in grade 10. I wasn't into it. I didn't put it in my long-term memory. It was a little tedious and hard for me to understand.

It seems that she may have confused a circuit breaker with some sort of switch used in a laboratory exercise on electrical circuits. More importantly her comment reveals a noticeable lack of interest in the topic.

On the other hand, only three out of the eight males interviewed used the school experience category for this item; the other five males referred to a range of informal experiences such as:

> I built a basement suite downstairs and I took stagecraft as well.

> I'm interested in electronics so I know about circuit breakers.

Items related to the wearing of seat belts (item 6), the stopping distances of cars traveling at different speeds (items 1 and 2), or identifying the type of substances that must be kept away from an open flame (item 17) yielded similar results in terms of the males choosing to draw on their informal and everyday experiences more than did the females. It would appear that on many of the items males had some advantage in answering them correctly by virtue of their ability to draw on a wider range of out-of-school experiences.

A second source of data used to address the issue of the role of the item context on student responses was the spontaneous, affective comments (which typically were negative) made by the female students as each question was discussed. Close to half of the items elicited an immediate negative response from more than one of the females interviewed, whereas only one male responded negatively and that was to only one item. Typical of the negative spontaneous female responses to an item is the following: Shown a picture of a Bunsen burner (item 7) and asked about the proper procedure for lighting it, several females made comments like:

> Bunsen burners scare me. I never did learn how to use them.

> Well, I got my partner to do it. I went and did the other stuff. Maybe I was afraid of Bunsen burners.

Similar negative comments were frequently voiced by the females as they were asked to explain their thinking on other items involving physical science concepts like circuit diagrams (item 19) or astronomy concepts (items 3 and 16):

> Oh no, not another physics item!
>
> I remember doing a thing on this but most of this comes from talking to the guys. I'm not into technical stuff.

These negative kinds of reactions may have blocked any serious cognitive engagement with the item and resulted in the females simply guessing at a response rather than trying to reason or even to recall some knowledge that might be of some assistance to them.

While our presentation of the results on the role of item context in creating (or at least exacerbating) gender differences has been brief, we think such results are indeed suggestive of an important factor that has frequently been overlooked in the construction of achievement tests.

Discussion

The data that we have sketched out above are consonant with other recent commentators in this field (Bateson & Parsons-chatman, 1989; Kelly, 1987a; Linn & Hyde, 1989; Sjoberg, 1989). Our results suggest that males seem to enjoy an advantage when it comes to retrieving knowledge or experiences that might be helpful in answering science achievement items. This advantage appears to be operative in two distinct but related ways. First, it seems that many of the items were embedded in contexts that allowed the males to utilize their prior everyday or "informal experiences"—contexts that typically did not seem to be accessible to the females in our study. Thus the males talked about their knowledge obtained from activities like wiring a basement suite, experiences with driving and fixing cars, using a compass in an orienteering course, or messing around with chemistry sets and model rockets. This is not to say that females do not have any opportunities to become involved in activities like this, but that on balance these sorts of activities tend to be more associated with the socialization experiences that are part of being brought up as a male in our society (Fischer, 1982; Johnson & Murphy, 1984; Lie & Bryhni, 1983). Furthermore, our interview results indicated that for items that involve topics like electricity, braking distances for cars, and appropriate safety procedures for working with chemicals, the females endeavored to draw on previous school experiences—experiences that often did not make much sense to them at the time when taught, let alone now. It is not surprising, then, that even though the females attempted to utilize their school knowledge on slightly more than half of the items (see Table 3), this strategy was not particularly successful for many of them and they still, on average, performed less well on these items than the males.

This finding, although different in origin, bears some similarity to the point made by Linn and Hyde (1989) about how the use of school-based knowledge and techniques may actually hinder female performance in mathematics achievement testing. They argue that teachers often emphasize the importance of going through all the steps of a mathematical algorithm or procedure and showing one's work. However, this is very time-consuming when solving a large number of problems on a timed achievement test. The end result is that those students who use these school-learned methods often do not complete the test. And the evidence seems to suggest that females are

more likely to resort to the techniques that they learned in school than are males.

It is interesting to note that in those situations where the females were able to make some connections to their life experiences, they were more successful in answering the questions. For example, on a question concerning the molecular motion of steam, a female explained how she arrived at her correct response by stating: "What I thought about steam is that it moves about a room. When you put a kettle on the stove and close the kitchen door the windows fog up." Or, in another question on the increase of water pressure with depth, several females indicated that they knew the answer because of their experience with diving. The more general point to be made, then, is that many of the items used to assess the understanding of physical science concepts confer an advantage to males because the item contexts enable them to deploy more effective answering strategies.

Another finding of general interest in our study was that a number of the items elicited spontaneous negative reactions by a number of the females. Many of the items referred to objects or situations where the females had frightening experiences (such as with a Bunsen burner), where they lacked confidence and experience (the electricity items), or which the females construed as boring and uninteresting ("not another physics item"). The result tended to be that the female students approached such items tentatively or with disinterest. On the other hand, they appeared to show more enthusiasm for items that made more intuitive sense to them.

These issues and results could be interpreted in the more narrow sense as suggesting we simply ensure that we develop assessment questions that are gender neutral. While we would certainly encourage such endeavors, we also agree with the critics who argue that the assessment issue is only a very small symptom of a much larger set of educational and sociocultural problems. We must ultimately reconceptualize our basic ideas about the nature of scientific inquiry (Easlea, 1981; Keller, 1985; Martin, 1988, 1989), and in particular our ideas about the design and delivery of science curriculum materials and instructional strategies (Bentley & Watts, 1986; Lewis & Davies, 1988; Manthorpe, 1982; McClintock Collective, 1988; Smail, 1987).

Our study points to some interesting directions for initiating some aspects of this reconceptualization process. One such direction entails a more careful examination of the subtle yet pervasive ways that prior experience and the accompanying role expectations constructed from these experiences influence both students and teachers in their daily interactions in the classroom. The importance of this influence of prior knowledge on subsequent learning activities has been the subject of intense scrutiny in the area of science and mathematics concept learning (Cobb, 1988; Driver, 1989; Novak, 1987; von Glasersfeld, 1989; White, 1988). The methods and findings from the extensive literature generated by this constructivist frame for educational inquiry would seem to have significant implications for extending the scope and shifting the nature of the research agenda in the field of gender and science (Vegting, 1987).

A brief example will illustrate what this shift in perspective would involve. Some researchers have argued that the number and kind of science courses

taken by males in comparison with females is an important factor in accounting for the differential performance in the physical sciences (Kahle, 1985; Kelly, 1981; Linn & Hyde, 1989). A more complex version of this argument has been advanced by Klainin and Fensham (1987) and Klainin, Fensham, and West (1989b) as a partial explanation of the results they obtained from a Thailand study where the females outperformed the males in several tests of chemistry achievement and did equally as well in a set of physics tests. The fact that the Thailand educational system required all students enrolled in the science curricular stream to study chemistry, physics, and biology at the secondary level is no doubt an important factor. However, given the implications that we have drawn from our study, we would suggest that their reporting of a different set of cultural norms and expectations regarding gender and science (e.g., the perception by the students that chemistry activities are "women's work" and that over 90% of the chemistry teachers in their study were female) represents a very provocative and interesting finding. These differing norms and attitudes seem to lead to behavioral outcomes on the part of the female students that has not been reported in the Western literature (e.g., the females in the chemistry classes were found to be "more dominant in classroom exchanges than the boys" (Klainin & Fensham, 1987).

Rather than attempting to quantify and classify school experiences in terms of the science courses taken, a constructivist perspective would focus more on an examination of the nature of the curricular content and the instructional strategies so as to determine the extent to which they acknowledge the prior experiences and understandings of all students. Further, it would be more concerned with the quality of the interactions between teachers and students and endeavor to illustrate the importance of continually checking out the meaning that students, as well as teachers, confer on classroom events and discourse. As a case in point, Erickson and MacKinnon (1991) recently documented in some detail an instance of a teacher who followed up on a simple comment made by three females working together on a laboratory problem involving electrical circuits. Initially the teacher thought that the students had the "correct" answer to the problem and that they understood the underlying concepts. But he was uneasy about one of their diagrams, and with some probing, combined with considerable patience and encouragement, the students discovered that their earlier understanding was quite confused and inadequate. Through additional experiments and considerable discussion and debate among themselves they were able to construct a more complete and functional understanding of electrical circuits. This short teaching encounter proved to be extremely satisfying for the students and, as a result, the teacher.

One of the important messages of these types of in-depth, classroom-based case studies is that they reveal to practitioners as well as researchers not only the inherent complexity of much of what we attempt to teach students but, more importantly, the value of listening carefully and checking out the meanings that students often attribute to their classroom experiences. This advice would appear to be particularly relevant when interacting with females, given the evidence presented above and the documented ten-

dencies of females to conform to school procedures and to produce "school answers" (Haggerty, 1987; Linn & Hyde, 1989).

In conclusion, we suggest that the present study represents a useful starting point for further explorations on gender and science. The qualitative data presented above raise a series of important and provocative questions. To what extent will females continue to rely on prior school experiences when in most instances this strategy did not appear to be fruitful in assisting them to clarify or answer assessment questions? Is this response pattern an example of females lacking confidence in their own personally constructed knowledge (Linn & Hyde, 1989), or is it a matter of the females being more sensitive to the implicit messages of school assessment exercises? That is, do females assume these exercises are measuring school learning and therefore attempt to provide appropriate school-based responses (Klainin, Fensham, & West, 1989b)? Is it possible to develop gender neutral assessment items as some researchers have encouraged? What is meant by gender neutrality as it pertains to achievement testing or curriculum development? Our results and the work of others (McLarty, Noble, & Huntley, 1988; Wedman & Stage, 1983) indicate that removing the gender bias in achievement testing needs to go well beyond the "superficial gender-related item modifications" tried by McLarty et al. (1988) to a more systematic examination of the experiential contexts of items.

References

Bateson, D., Anderson, J., McConnell, V., Rutherford, C., & Dale, T. (1986). *The 1986 British Columbia science assessment: General report*. Victoria, BC: Ministry of Education.

Bateson, D., & Parsons-Chatman, S. (1989). Sex-related differences in science achievement: A possible testing artifact. *International Journal of Science Education, 11*(4), 371-385.

Bearlin, M., Annice, C., & Elvin, A. (1988). *WASTE: Women and science education*. Canberra: Commonwealth Schools Commission.

Bentley, D., & Watts, M. (1986). Courting the positive virtues: A case for feminist science. *European Journal of Science Education, 8*, 121-134.

Crossman, M. (1987). Teachers' interactions with girls and boys in science lessons. In A. Kelly (Ed.), *Science for girls?* Milton Keynes, UK: Open University Press.

Cobb, P. (1988). The tension between theories of learning and theories of instruction in mathematics education. *Educational Psychologist, 23*, 87-104.

Daniels, J., & Kahle, J. (Eds.). (1987). *GASAT IV: Contributions from the fourth conference of Girls and Science and Technology*. Ann Arbor, MI: University of Michigan.

Driver, R. (1989). Changing conceptions. In P. Adey (Ed.), *Adolescent development and school science*. London: Falmer.

Easlea, B. (1981). *Science and sexual oppression*. London: Weidenfeld & Nicolson.

Erickson, F. (1986). Qualitative methods in research on teaching. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119-161). New York: Macmillan.

Erickson, G., & Erickson, L. (1984). Females and science achievement: Evidence, explanations, and implications. *Science Education, 68*(2), 63-89.

Erickson, G., & MacKinnon, A. (1991). Seeing classrooms in new ways: On becoming a science teacher. In D. Schön (Ed.), *The reflective turn: Case studies of reflection in and on practice*. New York: Teachers College Press.

Farkas, S. (1986). *Gender differences in science tests*. Unpublished master's thesis, University of British Columbia.

Finn, J., Dulberg, L., & Reis, J. (1979). Sex differences in educational attainment: A cross-national perspective. *Harvard Educational Review, 49*(4), 477-503.

Firestone, W. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. *Educational Researcher, 16*(7), 16-21.

Fischer, C. (1982). Science and the environment of young girls. In J. Ferguson (Ed.), *Who turns the wheel?* Ottawa: Science Council of Canada.

Haggerty, S. (1987). Gender and science achievement: A case study. *International Journal of Science Education, 9*(3), 271-279.
Harding, J. (Ed.). (1986). *GASAT III: Contributions from the third conference of Girls and Science and Technology*. London: Kings College, University of London.
Howe, K.R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher, 17*(8), 10-16.
Johnson, S., & Murphy, P. (1984). The underachievement of girls in physics: Toward explanations. *European Journal of Science Education, 6*(4), 399-409.
Kahle, J. (Ed.). (1985). *Women in science: A report from the field*. London: Falmer.
Kahle, J., & Lakes, M. (1983). The myth of equality in science classrooms. *Journal of Research in Science Teaching*, 20(2), 131-140.
Keller, E.F. (1985). *Reflections on gender and science*. New Haven: Yale University Press.
Kelly, A. (Ed.). (1981). *The missing half: Girls and science education*. Manchester, UK: Manchester University Press.
Kelly, A. (Ed.). (1987a). *Science for girls?* Milton Keynes, UK: Open University Press.
Kelly, A. (1987b). Does that train-set matter? Scientific hobbies and science achievement and choice. In J. Daniels & J. Kahle (Eds.), *GASAT IV* (Vol. 1). Ann Arbor, MI: University of Michigan.
Kimball, M. (1982). Sex differences in intellectual ability. In J. Ferguson (Ed.), *Who turns the Wheel?* Ottawa: Science Council of Canada.
Klainin, S., & Fensham, P. (1987). Learning achievements in upper secondary school chemistry in Thailand: Some remarkable sex reversals. *International Journal of Science Education, 9*, 217-227.
Klainin, S., Fensham, P., & West, L. (1989a). Successful achievements by girls in physics learning. *International Journal of Science Education, 11*, 101-112.
Klainin, S., Fensham, P., & West, L. (1989b). The superior achievement of girls in chemistry and physics in upper secondary schools in Thailand. *Research in Science & Technological Education, 7*(1), 5-14.
Koshland, D. (1988). Women in science. *Science, 239*, 1473.
Lewis, S., & Davies, A. (1988). *Gender equity in maths and science*. GAMAST Professional Development Manual. Canberra: Curriculum Development Centre.
Lie, S., & Bryhni, E. (1983). Girls and physics: Attitudes, experiences and underachievement. In S. Sjoberg (Ed.), *GASAT II*, Oslo: Science Education Center, University of Oslo.
Linn, M., & Hyde, J. (1989). Gender, mathematics and science. *Educational Researcher, 18*(8), 17-27.
Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences*. Stanford, Cal.: Stanford University Press.
Manthorpe, C. (1982). Men's science, women's science or science: Some issues related to the study of girls' science education. *Studies in Science Education, 9*, 65-80.
Martin, J.R. (1988). Science in a different style. *American Philosophical Quarterly, 25*, 129-140.
Martin, J.R. (1989). What should educators do about gender bias in science? In K. Tobin & M. Matthews (Eds.), *Proceedings of International Symposium on History and Philosophy of Science and Science Teaching*. Tallahassee, FL: Florida State University.
Marton, F. (1988). Phenomenography: Exploring different conceptions of reality. In D.M. Fetterman (Ed.), *Qualitative approaches to evaluation in education: The silent scientific revolution*. New York: Praeger.
McClintock Collective. (1988). *Getting into gear*. Melbourne: Ministry of Education.
McLarty, J., Noble, C., & Huntley, R. (1988, April). *Effects of item wording on sex bias*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans.
Novak, J. (Ed.). (1987). *Proceedings of the second international seminar: Misconceptions and educational strategies in science and mathematics*. Ithaca, NY: Cornell University.
Parsons-Chatman, S. (1987). Females and physical science: Is tinkering an issue? In J. Daniels & J. Kahle (Eds.), *GASAT IV* (Vol. 1). Ann Arbor, MI: University of Michigan.
Raat, J. (Ed.). (1981). *GASAT I: Contributions from the first conference of Girls and Science and Technology*. Eindhoven: Eindhoven University Press.
Sherman, J. (1978). *Sex-related cognitive differences: An essay on theory and evidence*. Springfield, IL: Thomas.

Sjoberg, S. (Ed.). (1983). *GASAT II: Contributions from the second conference of Girls and Science and Technology*. Oslo: Centre for Science Education, University of Oslo.

Sjoberg, S. (1989). Gender and the image of science. In P. Adey (Ed.), *Adolescent development and school science*. London: Falmer.

Smail, B. (1987). Organizing the curriculum to fit girls' interests. In A. Kelly (Ed.), *Science for girls?* Milton Keynes, UK: Open University Press.

Smith, J. K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher, 15*(1), 4-12.

Soltis, J. (1984). On the nature of educational research. *Educational Researcher, 13*(10), 5-10.

Taylor, H., Hunt, R., Sheppy, J., & Stronck, D. (1982). *The 1982 British Columbia science assessment: General report*. Victoria, BC: British Columbia Ministry of Education.

Vegting, P. (1987). Concept development in the natural sciences: Stimulating for girls. In J. Daniels & J. Kahle (Eds.), *GASAT IV* (Vol. 3). Ann Arbor, MI: University of Michigan.

von Glasersfeld, E. (1989). Cognition, construction of knowledge and teaching. *Synthese, 80*(1), 121-140.

Wedman, I., & Stage, C. (1983). The significance of contents for sex differences in test results. *Scandinavian Journal of Educational Research, 27*, 49-71.

White, R. (1988). *Learning science*. Oxford: Blackwell.

Education and Society is an established international journal which publishes original research and contributions dealing with major up-to-date issues concerning the relationship between education and society. It provides a new forum, where teachers and scholars throughout the world, working in education and related areas, are able to examine and evaluate current issues and problems in education and society — from a broad and balanced social, cultural and economic perspective. The journal, by focusing on both education and society, hopes to appeal to many educators and scholars from diverse educational backgrounds. Each issue of the journal contains a special symposium concerned with some of the more immediate problems affecting education, educational systems and society in a changing world.

The journal should be of interest to teachers, academics, research scholars, educational administrators and consultants, students in teacher training institutes and social workers.

Contributions are invited (normally no more than 5000 words, typed double spaced in two copies) and should be sent to the Editor, *Education and Society,* James Nicholas Publishers, P.O. Box 244, Albert Park, Australia, 3206.

In order to ensure that articles are of the highest quality, all contributions are submitted to at least two independent referees before acceptance for publication.

The journal appears twice a year, in June and December. These two issues constitute one volume. An annual index is included in the December issue.
1990, Volume 8, **ISSN 0726-2655.**

Subscription rates (payable in advance): one yearUS$90.00. Individuals and elementary and secondary schools US$48.00 (for accelerated surface postage and packing add US$5.00 per issue). Orders and remittances should be addressed to:

James Nicholas Publishers, P.O. Box 244, Albert Park, Australia, 3206.

EDITOR **Dr Joseph Zajda**

CONSULTING EDITORS
Professor Orlando Albornoz, *University of Venezuela*
Professor John Eggleston, *University of Warwick*
Professor Nigel Grant, *University of Glasgow*
Professor Henry Levin, *Stanford University*
Professor David Nyberg, *State University of New York (Buffalo)*
Professor Douglas Sloan, *Columbia University*
Professor Seth Spaulding, *International Bureau of Education, Geneva*
Dr. Joan Whitehead, *University of Cambridge*
Dr. John Wilson, *University of Oxford*
Professor Charles Willie, *Harvard University*

RECENT AND FORTHCOMING ISSUES
Ivan Illich **ON EDUCATION**
Martin Carnoy **EDUCATION AND THE THEORIES OF THE STATE**
Edmund King **CHANGING EDUCATION FOR A CHANGING SOCIETY**
John Eggleston **SCHOOLING AND WORK**
John Martin Rich **MERITOCRACY AND EQUAL EDUCATIONAL OPPORTUNITY**
James Bowen **SCIENCE, EDUCATION AND THE ENVIRONMENT**
Francis Bregha & Don Dawson **WORK, EMPLOYMENT AND LEISURE**
Stanley Aronowitz **WORKING CLASS IDENTITY AND CELLULOID FANTASIES**
Michael O'Loughlin & Malcolm Campbell **TEACHER PREPARATION, EMPOWERMENT AND REFLECTIVE INQUIRY**
John Smyth **TEACHERS-AS-INTELLECTUALS**
Joseph Zajda **EDUCATIONAL REFORMS AND IDEOLOGY IN SOVIET EDUCATION**
H. Svi Shapiro **EDUCATION AND THE CRISIS ON SOCIETY**
Nigel Grant **IMPACT OF COMPARATIVE EDUCATION ON SOCIETY**
Michael Grady & Charles Willie **COMMUNITY PARTICIPATION IN EDUCATION**
Gunnar Berg **THE SCHOOL ORGANIZATION**
Joseph Zajda **CURRICULUM REFORMS IN SOVIET SECONDARY SCHOOLS**

EDUCATION AND SOCIETY

ORDER FORM

☐ **Please enter our subscription to *Education and Society***

We enclose $

☐ **Please enter my subscription under the terms of the Personal Subscription Plan**
(please be sure to give your private address)

We enclose US$

☐ **Please charge my**
Access/American Express/Barclaycard/Diners Club/Mastercard/Visa account number:

Card expiry date **Signature**

Name

Address

.....

Signed **Date**

The Alberta Journal of Educational Research *Vol. XXXVII, No. 3, September 1991, 241-257*

BIKKAR S. RANDHAWA
University of Saskatchewan

Gender Differences in Academic Achievement: A Closer Look at Mathematics[1,2]

The present study addressed the issue of change in gender differences in academic achievement, more specifically in mathematics, using the effect size as a standard unit of comparison. The effect size was used both within and between years to discern gender differences on specific tests in mathematics and language areas administered to a provincial stratified cluster (classroom) sample of students at grades 4, 7, and 10. Additionally, mathematics and reading data collected in each of three consecutive years from a sample of Grade 10 students were analyzed to determine whether gender differences were consistent and systematic over three years. Results suggest that in grade 4, females increased their achievement advantage on Capitalization and Mathematics Concepts subtests between 1978 and 1985. The female achievement advantage observed in grade 7 in 1978 was significantly reduced by 1985 in Vocabulary, Capitalization, and Mathematics Concepts subtests. In grade 10, males outperformed females in mathematics while females achieved higher than males in language/reading. These grade 10 male/female advantages did not significantly change between 1978 and 1985. Furthermore, in mathematics components, the grade 10 results from one school were consistent and systematic, indicating consistent male advantage in computation, concepts, and problem solving. These differences cannot be attributed to differences in perceived classroom behaviors of males and females by their teachers nor to differences in cognitive structures.

Gender differences in cognitive tasks have intrigued researchers since the turn of this century (Jacklin, 1989). However, since the publication of the Benbow and Stanley (1980) study on mathematically precocious youth, there has been considerable controversy regarding both the extent and the reasons for the gender differences in achievement, more specifically in mathematics involving measurement using standardized tests (Hyde, Fennema, & Lamon, 1990; Randhawa & Hunt, 1987) and using teacher assigned grades (Kimball, 1989). The focus of the present article is on gender differences in mathematics, but it also deals with differences in language related subtests.

Gender Differences

Mathematics

Generally, until about grade 4 either no differences between males and females are observed in mathematics achievement (Callahan & Clements, 1984; Dossey, Mullis, Lindquist, & Chambers, 1988; McKay, 1979; Randhawa & Hunt, 1987) or differences that are observed favor females (Brandon, Newton, & Hammond, 1985; Hawn, Ellet, & Des Jardines, 1981;

Bikkar Randhawa is a professor of educational psychology with current research interests in cognitive psychology and gender-related differences in mathematics learning and instruction.

Randhawa & Hunt, 1987). An exception to these findings was reported by Lewis and Hoover (1986). They observed that on the standardization of the Iowa Tests of Basic Skills, grade 2 boys achieved at a higher level than girls on the problem solving subtest but that this difference decreased in grade 5 and disappeared by grade 8.

Mixed results have been reported for the junior high school grades (years). Either small gender differences in favor of boys (Hilton & Berglund, 1974; Raphael, Wahlstrom, & McLean, 1984; Tsai & Walberg, 1979) or no differences between males and females (Cicirelli, 1967; Connor & Serbin, 1985; Raphael et al., 1984; Randhawa & Hunt, 1987) have been reported. For example, the results of the Ontario portion of the Second International Mathematics Study (Raphael et al., 1984) involving grade 8 students indicated that boys scored significantly higher than girls on 19 items, girls scored significantly higher than boys on 10 items, and there were no differences on 11 items. Moreover, it was found that boys tended to perform better on the more difficult questions. Exceptions to these mixed results are studies of "gifted" youth in which consistent reliable differences in favor of boys were obtained (Benbow & Stanley, 1982; Weiner, 1984). However, it must be kept in mind that in these latter studies gender differences were examined in terms of total SAT scores rather than in terms of item (Raphael et al., 1984) or component score (Randhawa & Hunt, 1987) differences.

On overall math achievement scores throughout the high school years, differences favoring boys are common (Jones, Burton, & Davenport, 1984; Ramist & Arbeiter, 1986; Randhawa, 1988; Randhawa & Hunt, 1987). However, gender differences favoring males are more commonly and consistently found on problem solving components and applications (Carpenter, Lindquist, Matthews, & Silver, 1983; Randhawa, 1988; Randhawa & Hunt, 1987; Swafford, 1980).

Verbal and Language Arts

Gender differences in mathematics achievement favoring males, especially at the high school level and beyond, have received considerable attention and debate (Wittig & Petersen, 1979). However, the consistent female advantage in verbal tests and skills generally throughout the school years (Alberta Department of Education, 1989; Green, 1987; Myerberg, 1986; Randhawa & Hunt, 1987) has not been the source of much debate and concern regarding the causative factors explaining female superiority. Even beyond the high school, consistent but declining female advantage in verbal tests is reported (Wilder & Powell, 1989).

Wilder and Powell (1989) provide an extensive review of the significant research on gender differences in test performance, including studies documenting these differences and theories put forward for the factors related to differential performance. This review includes studies involving undergraduate admission tests, graduate and professional school admission tests, validity studies, national studies, verbal ability tests, quantitative ability tests, and voluntary testing programs. Explanations of test performance differences reported in the literature, according to this survey, include biological, social, psychological, demographic, and educational factors as well

as test characteristics. Some of these explanations are discussed in more detail in the next section in relation to gender achievement differences in mathematics.

Wilder and Powell (1989) conclude that there is a convergence in the literature on gender differences in that males have caught up with females on tests of verbal ability and achievement, while females have narrowed the gap between themselves and their male counterparts on tests of mathematics achievement. Whether similar trends emerge in a provincewide data on language/verbal tests and mathematics tests is one of the objectives of this article. Also, this article explores the underlying causes, albeit empirically and speculatively, of gender differences in achievement in both of these areas.

Explanations of Gender Differences

Psychologists have attempted to explain gender differences in mathematics achievement on the basis of underlying ability (Benbow & Stanley, 1980; 1981; 1982; 1983) and differential course taking (Fennema & Sherman, 1977; Linn & Hyde, 1989). Although Benbow and Stanley speculate about biological causes for their observed gender differences, they had no data to substantiate their conclusion. Also, their studies did not take into account their subjects' differential math-related experiences and assumed that enrollment in equal numbers of mathematics classes would account for this factor (Eccles & Jacobs, 1986; Lips, 1988).

Stylistic information processing, as an alternative explanation to gender differences, was proposed by Dinnel and Glover (1987). Assumed in this explanation are the well-known concepts of storage and retrieval capacity and strategy, knowledge representation, planning, and execution. On the other hand, Fennema and Peterson (1985) suggest that gender-related differences in mathematics might be explained on the basis of the autonomous learning behavior of boys and girls. Choosing to do high-level tasks, working independently on these tasks, persisting in them, and attaining success on them are examples of autonomous learning behaviors. Fennema and Peterson suggest that males have more opportunities than females to engage in these behaviors both in the school and outside. Accordingly, the nature and type of contact between the teacher and the student determine the opportunity for the student for autonomous behaviors. They report that some types of contact enhance autonomous behaviors, but females are not as likely as males to have those contacts. In the same vein, Eccles and Jacobs (1986) attribute social forces as possible determiners of mathematics performance and attitudes.

On the basis of a meta-analysis of 98 studies of gender differences in mathematical tasks reported between 1975 and 1988, Friedman (1989) conclude that "we cannot say with 95% certainty that a gender difference exists in the general United States population of school-age youth" (p. 204). These studies included first graders to twelfth graders. However, a subset of 36 of these studies, in which the performance of high school-age students was examined, indicated that the gender difference in favor of males is diminishing as a function of the year of study (Friedman, 1989). She proposes an

environmental explanation for the decreasing gender differences in mathematics.

Huston and Carpenter (1985) speculate that general gender differences may in fact be present by the time children enter elementary schools. They found from their research with preschool children that girls spend more time in highly structured activities than boys. Structure in this case was determined by (a) the amount of feedback given by adults, and (b) the extent to which the child could observe adults performing activities similar to those in which the child was involved. Huston and Carpenter found that preschool children who were engaged in highly structured activities were more likely to seek approval from adults than those engaged in less structured activities. Therefore, it is probable that gender differences in responsiveness to adults that appear during the preschool years continue into the later years. Hence this tendency of differential responsiveness of males and females may stimulate teachers to react differently to males than to female students.

This article integrates the results from three research studies of gender differences in language and mathematics. The first two, in 1978 and 1985, were based on two provincewide assessments conducted at grades 4, 7, and 10. The third study, in 1987-1989, was based on a three-year assessment of reading and mathematics of grade 10 cohorts in a high school. More specifically, the questions addressed were:

1. What are the patterns of change in the effect sizes for various grades and tests between test occasions? Has the gap between the achievement of males and females in language and mathematics narrowed between 1978 and 1985 for grades 4 and 7 and between 1978, 1985, and 1987-1989 for grade 10?
2. For the grade 10 subjects tested in 1987, 1988, and 1989, are the gender differences in reading and mathematics consistent?
3. Are the classroom behaviors of males and females rated similarly by the students' teachers?
4. How do the factor structures of males and females on reading and mathematics microcomponents compare?

Method

Samples

1978 and 1985 studies. Similar sampling procedures were used for both the 1978 and 1985 studies. A single-stage stratified (by school district) sampling design was used to sample classes from each school jurisdiction in Saskatchewan so that 10% of the classes during the academic years 1978-1979 and 1985-1986 were included. Jurisdictions with fewer than six classes at each of the selected grade levels, 4, 7, and 10, were placed in a separate stratum and sampled at the 10% rate at each grade level. In all cases, all students in each sample class were tested. However, data for a few classes (between 1% and 2% of the classes) were not available due to failure to conduct the testing according to instructions or lost in the mail.

1987-1989 study. A high school in a major city in Saskatchewan was the source of data for a three-year study of reading and mathematics achievement. All those grade 10 students in each of the three years who were present

during the two days of test administration in mid-May were tested. This school serves a fairly large enrollment area and students from all socioeconomic backgrounds attend it. The distribution of students by gender and grade for each study is summarized in Table 1.

Tests

1978 study. Students in grades 4 and 7 in the 1978 study were administered the Canadian Tests of Basic Skills (CTBS), Form 4M (King, 1977). The grade 10 students were administered the Sequential Tests of Educational Progress (STEP), Series 2, Form 2A (Educational Testing Service, 1971). The specific subtests of the CTBS and STEP are listed in Table 2 in the order in which they were administered.

1985 study. The tests used for the 1985 study were the appropriate levels of the CTBS, Form 5M (King, 1982) for grades 4,7, and 10. The specific subtests of the CTBS for each grade and the order of administration are given in Table 2.

1987-1989 study. As shown in Table 2, Reading and Mathematics subtests of the CTBS, Form 5 were administered in the 1987-1989 study.

As well, teachers of Algebra 10 (a grade 10 course) were asked to rate selected students from their classes using the Florida Key (FK) (Purkey, Cage, & Graves, 1973) at the end of school year in June 1989. Altogether, 86 completed FKs were returned. Of these, 43 were for males and 43 for females.

The FK scale contains four subscales: Relating (5 items), Asserting (7 items), Investing (2 items), and Coping (4 items). Each item is rated using a 4-point Likert scale. As a research instrument, the reliability and validity data on this instrument are modest at best. Based on the nominations of 27 teachers of five students who the teachers thought felt "best about themselves as learners" and another five students who the teachers thought "felt badly about themselves as learners," point-biserial correlation coefficients for items with subscales ranged from .57 to .71. An interreliability of .84 for

Table 1
Distribution of Students by Sex, Grade, and Year of Study[a]

Year	Grade					
	4		7		10	
	M	F	M	F	M	F
1978	675	654	790	706	859	900
1985	915	832	837	849	834	753
1987					58	64
1988					58	72
1989					82	79

[a]The number of males and of females included in the 1978 and 1985 studies represented between 9.7 to 10.6% of the provincial enrollment by gender on September 30 in the respective grades.

Table 2
Standardized Tests Administered to the Various Samples and the Order of Administration of the Subtests

Grades 4 and 7		Grade 10		
1978	1985	1978	1985	1987-1989
CTBS[1], Form 4M	CTBS, Form 5M	STEP[2], Form 2A	CTBS, Form 5M	CTBS, Form 5M
Vocabulary	Reading	Reading	Reading	Reading
Reading	Mechanics of Writing	Mathematics	Mathematics	Mathematics
Language:		Written Expression		
Spelling	English Expression			
Capitalization		Using Sources		
Punctuation	*Mathematics:*	of Information		
Usage	Basic Concepts			
Mathematics:	Computation			
Concepts				
Problem Solving				
Computation[a]				

[1]CTBS is the Canadian Tests of Basic Skills.
[2]STEP is the Sequential Tests of Educational Progress.
[a]This subtest was included in Form 5M as a separate entity. Priorly, items on computation were included in the Problem Solving subtest.

the composite scale was reported using the data of three teachers who had rated the same 11 students.

Procedure

The testing for the 1978 and 1985 studies was completed during October and November. The tests were administered by the classroom teachers according to a schedule provided by the researcher (see Table 2). For the 1987-1989 studies the tests were administered in mid-May of each year. As shown in Table 2, the Reading test was administered first followed by the Mathematics test in each of the three years.

In June, 1989 the teachers provided data for the selected students on the FK.

Analyses

To answer the first question, the basic statistic calculated was the effect size. This is based on Hyde (1981) and corrected by the factor shown by Hedges (1981) to produce an unbiased estimator of the population effect size for both large and small samples. The corrected effect size for a test is given by

$$d = c\,(Y_F - Y_M)/s,$$

where d is the corrected effect size for a test in a given year, c is the correction factor, Y_M and Y_F are the means of males and females, respectively, and s is

the pooled standard deviation of both groups on the same test. The correction factor c is defined as

$$c = 1 - \frac{3}{4n-9},$$

where n is the total number of males and females. The estimated variance of d is found by

$$s^2(d) = \frac{n}{n_F n_M} + \frac{d^2}{2n}$$ (Hedges & Olkin, 1985, p. 86).

The 1978 and 1985 studies involved fairly large samples of over 1,300 subjects and the combined sample size of the 1987-1989 study was over 400 (see Table 1). With samples as large as these, the variance of the estimate of effect size becomes quite small and hence the estimate is very precise (Friedman, 1989). Therefore, small changes in effect sizes would indicate significant shifts in gender differences in the tested achievement. Furthermore, regardless of the effect size, the estimate of the standard error of an effect size is influenced markedly by the sample size of each group. Hence, as will be noted in the results, the standard errors of effect sizes for various subtests for the same two groups are almost identical.

For the second question, two separate 2 x 3 (gender-by-year) fixed effects multivariate analysis of variance (MANOVA) were performed. In the first, the dependent variables were the CTBS Reading and Mathematics total scores. In the second, the three mathematics component scores—computation (9 items), concepts (20 items), and problem solving (19 items) were the dependent variables. Not normally reported, derivation of these component subscores was based on the classification system provided in the CTBS manual. A similar MANOVA involving the components of the Reading test was not performed because, unlike the Mathematics test, none of the univariate effects for it were significant.

Given nonorthogonality due to unequal sample sizes, the analyses were conducted so that the effects other than the one of interest are first partialled out. This yields a more conservative estimate of the *F*-ratio than if this partialling had not been done.

In order to answer the third question, a MANOVA was performed on the four FK subscales following the same analysis approach as that used for the mathematics components.

The fourth question was addressed by the analyses described below. The CTBS, Level 16, Mathematics and Reading items were grouped according to the process and content classifications as described in the test manual. On each of these 21 microcomponents the score of each subject was computed. The data of male and female students from the high school, combined over three years, were processed separately. First a principal components factor analysis was performed. In each case two principal components were retained. This decision was based on the Cattell's (1966) scree test for the number of factors. The obtained principal factor structures were then subjected to an orthogonal Varimax rotation (Kaiser, 1958). Oblique rotation might be theoretically interesting for the achievement microcomponents, but for the purpose of the present study orthogonal rotation served the function

well. Furthermore, examination of the factor structures from both oblique and orthogonal rotations indicated that there was little difference in the simple structures or the factor interpretations. Lastly, these factor matrices were rotated to a procrustes congruence criterion (Schonemann, 1966) in order to determine whether the factor structures of males and females were invariant (congruent). The females' factor matrix was used as the target for this rotation.

Results

Question 1: What are the patterns of change in the effect sizes for various grades and tests between test occasions? Has the gap between the achievement of males and females in language and mathematics narrowed between 1978 and 1985 for grades 4 and 7 and between 1978, 1985, and 1987-1989 for grade 10?

Grade 4. Effect sizes for gender differences on the CTBS tests for the 1978 and 1985 samples of grade 4 students are given in Table 3. In 1978 females had higher observed means than males on all subtests except the Mathematics Concepts and Problems subtests ($p < .05$). Likewise, except for the Mathematics Concepts, grade 4 female students in 1985 outperformed the grade 4 males ($p < .05$).

The estimation of effect sizes is used in quantitative research syntheses of studies. However, in the present study, these estimated effect sizes are used to determine whether in the time span between the studies the female/male advantage has increased or decreased. Given the standard errors for various subtests in Table 3 for 1978 and 1985, the standard error of the difference between effect sizes for any subtest is .073. This is because the standard error

Table 3
Within Group Standard Deviations, Effect Sizes, and the Standard Deviations of the Effects for Grade 4 Samples

Subtest	SW		Female-Male Effect Size (d)				$d_{1985-1978}$
	1978	1985	1978		1985		
1. Vocabulary	6.316	7.267	.142*	(.055)[a]	.156*	(.048)	.014[b]
2. Reading	11.755	9.081	.103*	(.055)	.210*	(.048)	.107
3. Spelling	7.492	6.900	.355*	(.055)	.426*	(.048)	.071
4. Capitalization	6.613	5.444	.277*	(.055)	.436*	(.048)	.159*
5. Punctuation	7.010	5.494	.335*	(.055)	.351*	(.048)	.016
6. Usage	6.045	5.766	.251*	(.055)	.387*	(.048)	.136
7. Math Concepts	6.135	5.041	−.077	(.055)	.079	(.048)	.156*
8. Math Problems	4.998	5.103	.024	(.055)	.123*	(.048)	.099
9. Math Computation[c]	-	7.858	-	-	.285*	(.048)	-

*$p < .05$.
[a]The value in parentheses is the standard error of the effect size.
[b] 95% confidence interval for the difference is +.143.
[c]This subtest was included in Form 5M as a separate entity. Priorly, items on computation were included in the Problem Solving subtest.

of the effect size and the sample sizes for each subtest for a testing occasion were equal. Therefore, a change in the effect size of +.143 would lead us to conclude that the gap has either increased or decreased ($p < .05$). Using this criterion, it is concluded that the grade 4 average female student has gained further ground over the average male student on the Capitalization subtest between the two testing occasions. Similarly, on the Mathematics Concepts subtest the change between 1978 and 1985 was significant, but it might be attributed to chance because the difference between males and females on each testing occasion was not significant.

Grade 7. Effect sizes for gender differences on the CTBS subtests for the 1978 and 1985 studies are given in Table 4. In 1978, on all except the two mathematics subtests females' means were significantly higher than those of males ($p < .05$).

In 1985, positive and significant effect sizes were obtained on all subtests except Vocabulary, Mathematics Concepts, and Mathematics Problems ($p < .05$). In contrast, on Vocabulary the effect size was negative and significant, suggesting that on this subtest the males did better than the females. As with the grade 4s, the effect sizes for Mathematics Concepts and Problems were not significant. Interestingly, on the Mathematics Computation subtest, which was a part of the Problems subtest in the CTBS battery used in 1978, females outperformed males. Based on the criterion shift of + .14 ($p < .05$) in the effect size, it is evident that females lost ground to males on Vocabulary and Capitalization subtests. Similar to the results of the grade 4 studies on the Mathematics Concepts subtest, the effect size for each testing occasion was not significant. But the difference in the effect sizes between

Table 4
Within Group Standard Deviations, Effect Sizes, and the Standard Deviations of the Effects for Grade 7 Samples

Subtest	Sw		Female-Male Effect Size (d)		$d_{1985-1978}$
	1978	1985	1978	1985	
1. Vocabulary	8.071	8.136	.129* (.052)[a]	−.114* (.049)	−.243*[b]
2. Reading	13.307	9.800	.255* (.052)	.120* (.049)	−.135
3. Spelling	8.990	8.590	.489* (.053)	.471* (.049)	−.018
4. Capitalization	7.208	5.411	.497* (.053)	.340* (.049)	−.157*
5. Punctuation	7.327	5.994	.495* (.053)	.388* (.049)	−.107
6. Usage	5.878	5.264	.416* (.052)	.315* (.049)	−.101
7. Math Concepts	7.940	7.349	.088 (.052)	−.060 (.049)	−.148*
8. Math Problems	4.838	6.017	.012 (.052)	.025 (.049)	.013
9. Math Computation[c]	-	7.411	- -	.229* (.049)	-

$p < .05$.
[a]The value in parentheses is the standard error of the effect size.
[b] 95% confidence interval for the difference is +.143.
[c]This subtest was included in Form 5M as a separate entity. Priorly, items on computation were included in the Problem Solving subtest.

the two occasions on this subtest was significant, which might suggest that females lost ground to males between 1978 and 1985. It might be possible that this result is a chance artifact.

Grade 10. Effect sizes for gender differences on the STEP and CTBS subtests are provided in Table 5 for the three grade 10 samples. Females outperformed males on all four language-related subtests of the STEP ($p<.05$). On the other hand, males outperformed females on the Math Concepts subtest ($p < .05$). On the Math Computation subtest the difference was not statistically significant.

On the CTBS subtests in 1985 all four effect sizes were significantly different from zero ($p < .05$). Moreover, on the Reading, Written Expression, and Using Sources of Information (USI) subtests, females outperformed males, while on the Mathematics subtest males outperformed females. However, for the 1987-1989 sample, the effect size for the Reading subtest was not significantly different from zero. But a significant effect size on the Mathematics subtest with a negative sign, as was the case for the 1985 sample, was observed. Hence the average male outperformed the average female on this subtest.

To compare the results of 1978 and 1985, the average effect sizes across the mathematics subtests and across the language subtests within each year

Table 5
Within Group Standard Deviations, Effect Sizes, and the Standard Deviations of the Effects for Grade 10 Samples

Subtest	SW			Female-Male Effect Size (d)				
	1978	1985	1987-1989	1978	1985		1987-1989	
STEP								
Reading	8.52	-	-	.136* (.048)	-		-	
Spelling	7.70	-	-	.440* (.048)	-		-	
Capitalization	7.00	-	-	.493* (.048)	-		-	
English Expression	8.95	-	-	.331* (.048)	-		-	
Math Computation	9.68	-	-	.032 (.048)	-		-	
Math Concepts	6.30	-	-	−.175* (.048)	-		-	
CTBS								
Reading	-	9.28	9.45	-	.288*	(.051)[a]	.051	(.099)
Mathematics	-	7.98	8.48	-	−.116*	(.050)	−.330*	(.099)
Written Expression	-	8.74	-	-	.546*	(.051)		
Using Sources of Information	-	10.12	-	-	.259*	(.050)	-	
d_{math}				−.071	−.116		−.330	
$d_{language}$				.350	.364		.051	

$p < .05$.
[a] Standard error of the effect size.

were computed and compared. In the case of mathematics, the effect sizes were, respectively, –.071 and –.116. Since the difference between the two was less than .1 in absolute terms, it is concluded that the relative standing of males and females in mathematics did not change. The same conclusion is reached when we compare the average effect sizes of language. The average effect sizes for 1978 and 1985 on the language subtests were, respectively, .350 and .364. Moreover, if the USI subtest, which includes a number of mathematics items, is excluded, the average effect size for 1985 in language-related subtests is .417. The above conclusion is still valid.

Now, if the 1987-1989 results from the single school included in the 1987-1989 study are compared with the average effect sizes computed for the 1978 and 1985 samples, it might be concluded that, at the grade 10 level, females lost ground to males in both language and mathematics. However, such a comparison is not warranted given the differences in sample selection noted earlier. Thus it appears that there is no discernible shift in the achievement patterns of a provincially representative sample of males and females in mathematics and language-related areas between 1978 and 1985.

Question 2: For the grade 10 subjects tested in 1987, 1988, and 1989, are the gender differences in reading and mathematics consistent?

In the first analysis using Reading and Mathematics as dependent variables, the only significant effect was the Gender main effect ($F(2, 406) = 13.53$, $p<.05$). The corresponding univariate analysis revealed that the Gender effect was significant ($F(1, 407) = 11.69, p < .05$) only on Mathematics. The corresponding means for males and females were 30.20 and 27.39, respectively. Since the first-order interaction was not significant, it is evident that the males' advantage in achievement in mathematics over the females' was consistent and systematic over the three years.

The Gender main effect was also significant in the MANOVA analysis of the three mathematics component scores ($F(3, 405) = 4.64, p<.05$). Furthermore, the univariate Gender main effects (df = 1, 407) on Computation ($F=8.56$, $p<.05$), Concepts ($F=10.95$, $p<.05$), and Problem Solving ($F=12.08$, $p<.05$) were reliable. Again, the first order interaction between Gender and Year was not significant. On each of the three mathematics components, males' mean achievement was higher than that of females' in a consistent and systematic fashion over the three years considered (see Table 6).

Question 3: Are the classroom behaviors of males and females rated similarly by their teachers?

Means and standard deviations of males and females for the subscales of FK are given in Table 7. The MANOVA results indicated that the Gender effect was not significant. Hence it appears that the classroom behaviors of males and females were rated similarly by their teachers.

Question 4: How do the factor structures of males and females on reading and mathematics microcomponents compare?

Application of the scree test to each of the principal component solutions revealed two factors for both males and females. The congruence matrices of females and males and the commonalities of the variables are given in Table

Table 6
Means and Standard Deviations on the Mathematics Components of Grade 10 Students

Component	Year					
	1		2		3	
Males n	58		58		82	
Computation	6.1	(1.9)[a]	6.1	(2.0)	6.0	(2.1)
Concepts	12.1	(4.3)	13.2	(4.2)	12.6	(4.4)
Problem Solving	11.1	(3.8)	12.3	(3.4)	11.4	(4.2)
Females n	64		72		79	
Computation	5.3	(1.6)	5.6	(2.0)	5.7	(1.5)
Concepts	10.4	(3.6)	12.4	(3.8)	11.2	(3.4)
Problem Solving	9.7	(3.1)	10.9	(3.6)	10.4	(3.3)

[a]Standard deviation.

8. As indicated by the Tucker congruence coefficients for each pair of factors, .97 and .98 respectively, the two pairs of factors were highly congruent. This suggests that, in general, the operative common structures of males and females for handling language and mathematics microcomponents are quite similar. But the two principal factors capture 41.6% and 46.8% of the total variance in the set of 21 variables for females and males respectively. Hence there is enough idiosyncratic variance that could produce differences in achievement on these two tests.

The first factor in each case can be interpreted to be the language and the second the mathematics factor. However, the language factor for both females and for males accounts, respectively, for 81.9% and 84.4% of the common variance in the set of variables. Moreover, the two factor structures are not good simple structures because in each set of variables a number load on both factors with a loading greater than .30. Taken the above considerations into account and that of parsimony, it may well be that there is only a single common factor in each set and which could parsimoniously be labeled as the achievement factor. This is not to deny that there may be specific cognitive components which are more instrumental for doing well on certain

Table 7
Means and Standard Deviations of Males and Females on FK Subscales

Subscale	Male (n = 43)		Female (n = 43)	
Relating	20.2	(2.6)[a]	20.6	(3.0)
Asserting	23.4	(7.8)	24.5	(4.5)
Investing	5.6	(2.6)	5.4	(2.2)
Coping	14.4	(3.5)	14.0	(3.9)

[a]Standard deviation.

Table 8
Congruence Factor Matrices of Males and Females

Variable	Females			Males		
	h^2	I	II	h^2	I	II
Textbook Factual Material	69[a]	78	28	66	75	31
Textbook Inferential Material	66	74	32	73	82	22
Textbook General Material	45	67	03	53	72	13
Everyday Factual Material	59	72	26	66	75	32
Everyday Inferential Material	66	72	38	66	76	29
Everyday General Material	58	73	21	52	65	32
Operational Computation	45	32	59	52	36	63
Equivalent Forms and Order Concepts	50	24	67	62	34	71
Equivalent Forms and Order Problem Solving	17	–04	41	39	–17	60
Common Applications Computation	16	02	41	36	08	60
Common Applications Concepts	40	26	57	36	26	54
Common Applications Problem Solving	55	43	61	66	41	70
Algebra Computation	35	44	40	23	14	46
Algebra Concepts	19	05	44	36	31	52
Algebra Problem Solving	29	27	47	18	24	35
Geometry and Measurement Computation	19	37	23	20	27	35
Geometry and Measurement Concepts	28	21	49	43	39	53
Geometry and Measurement Problem Solving	44	11	65	45	20	64
Statistics, Graphs, and Table Concepts	38	42	45	44	38	54
Statistics, Graphs, and Table Problem Solving	26	31	40	40	42	46
Basic Math Problems Concepts	50	38	60	50	40	58
Eigenvalue		7.17	1.58		8.30	1.54
Total Commonality	8.75					
Tucker Coefficient		97	.88			
Common Variance		81.9	18.1		84.3	15.7

[a]Leading decimal places are omitted.

cognitive tasks than on others, but that these various components are united by a common cognitive orientation. It would appear from our evidence that males and females are not differentiated on cognitive orientation.

Discussion and Conclusions

The studies reported here have examined gender-related differences in language and mathematics in grades 4, 7, and 10. An interesting picture emerges. At the grade 4 level in 1978, girls outperformed boys in the language and verbal subtests but there were no differences on the mathematics subtests. However, in 1985, females outperformed males on all except the Mathematics Concepts subtest. Comparison of the gender effect sizes for grade 4 students in 1978 and 1985 indicated that females had increased their ad-

vantage over males on the Capitalization subtest. This finding is not in conformity with the conclusion of Friedman (1989), based on a meta-analysis, and of Wilder and Powell (1989), derived from a review of the literature. They concluded that the general gap favoring males in mathematics achievement is narrowing with the passage of time. Moreover, Wilder and Powell (1989) concluded that males have caught up with females in tests of verbal ability and achievement. This does not appear to be the case with grade 4 students in language and verbal subtests. Additionally, females either outperformed males on the mathematics subtests or there were no differences.

The opposite of this trend and consistent with Wilder and Power's (1989) conclusion was the case for our grade 7 cohort studies. Females were significantly higher than males on all the language-related subtests in 1978 but on the two mathematics subtests there were no significant differences. However, in 1985, females were still significantly higher than males on all the language-related subtests except Vocabulary, on which males were significantly higher than females. Furthermore, females were significantly higher than males on Mathematics Computation subtest in 1985. Again, contrary to Friedman's (1989) conclusion, but in the opposite direction to what was observed in grade 4, females had lost ground in their advantage over males on Vocabulary and Capitalization subtests. Furthermore, on the Mathematics Concepts subtest, even though the effect sizes in 1978 and 1985 were not significant, the difference in the effect sizes was significant. This might suggest that females had lost ground on Mathematics Concepts subtest in 1985. But caution should be exercised with this interpretation because the result might be a chance artifact.

Consistency with Friedman's (1989) and Wilder and Powell's (1989) conclusion of gender differences in achievement was found in grade 10 cohort studies. Males outperformed females in mathematics and females outperformed males on language-related subtests in 1978 and 1985. In the case of the male advantage in mathematics, the results of the grade 10 study, which included students from only one high school, indicated that the male advantage was consistent and systematic during the three-year period considered.

In spite of the narrowing gap in achievement in language-related areas observed in grade 7, the gap in this area still favors females at grade 10. But the reversal in mathematics achievement differences at grade 4, mostly favoring females, has occurred at the grade 10 level. The gap has narrowed at grade 7 but by grade 10, consistent and systematic differences in all three mathematics components, Computation, Concepts, and Problem Solving, were observed. Preliminary results from a study of mathematics achievement at grade 12 and postsecondary levels seems to confirm this systematic and consistent gender difference in favor of males (Randhawa & Beamer, in progress; Randhawa, Beamer, & Ng, in progress).

How do we explain the differences? At the grade 10 level, we attempted to determine if teachers' ratings of males' and females' classroom behaviors would provide a clue. It seems that teachers found coping, investing, asserting, and relating behaviors of males and females quite similar. But caution is in order because the data set was small and the instrument does not appear

to have satisfactory reliability. Also, the attempt to determine if different cognitive structures of males and females contributed to processing differences that in turn might have led to observed differences did not bear differential evidence. The fact that the two common factors turned out to be invariant suggests that cognitive processing structures of males and females are similar. Because the common factors account for only less than half of the variance, an argument can be made that there might be other idiosyncratic mechanisms of males and females that account for systematic achievement differences. These idiosyncratic mechanisms could include the autonomous learning behavior (Fennema & Peterson, 1985) or the social forces and motivations (Eccles & Jacobs, 1986). Further systematic research of classroom processes and individual learner characteristics might shed further light on the sources of the gender differences. In the meantime, instructional processes must provide for equitable access and opportunity for both males and females to profit from them optimally.

Notes

1. This paper was prepared while the author had a research time stipend from the Social Sciences and Humanities Research Council of Canada. This support is gratefully acknowledged. However, the opinions in and substance of the paper are the sole responsibility of the author and do not imply an endorsement by the funding agency.
2. Sheryl Yuzik, Jason Randhawa, and Lisa Randhawa assisted in the preparation of this paper. Their contribution is much appreciated. However, I am fully responsible for any errors and conceptual fallacies and problems that may still exist in this paper. Also, the cooperation of the school administration and teachers is gratefully acknowledged. We are grateful to the students who supplied the data and without whose participation this paper would have been nothing but a dream. Helpful comments of an anonymous reviewer and of Todd Rogers, the guest editor of this section, are gratefully acknowledged. Send requests for reprints to Dr. Bikkar S. Randhawa, Department of Educational Psychology, College of Education, University of Saskatchewan, Saskatoon, Canada S7N OWO.

References

Alberta Department of Education. (1989). *Achievement testing program. Provincial report: June 1989*. Edmonton, AB: Author. (ERIC Document Reproduction Service No. ED 316577)

Benbow, C.P., & Stanley, J.C. (1980). Sex differences in mathematics ability: Fact or artifact? *Science, 210*, 1262-1264.

Benbow, C.P., & Stanley, J.C. (1981). Mathematical ability: Is sex a factor? *Science, 212*, 118, 121.

Benbow, C.P., & Stanley, J.C. (1982). Consequences in high school and college of sex differences in mathematical reasoning ability: A longitudinal perspective. *American Educational Research Journal, 19*, 598-622.

Benbow, C.P., & Stanley, J.C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science, 222*, 1029-1031.

Brandon, P.R., Newton, B.J., & Hammond, O.W. (1985, April). *The superiority of girls over boys in mathematics achievement in Hawaii*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Callahan, L.G., & Clements, D.H. (1984). Sex differences in rote-counting ability on entry to first grade: Some observations. *Journal for Research in Mathematics Education, 19*, 598-622.

Carpenter, T.P., Lindquist, M.M., Matthews, W., & Silver, E.A. (1983). Results of the third NAEP mathematics assessment: Secondary school. *Mathematics Teacher, 76*, 652-659.

Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Cicirelli, V.G. (1967). Sibling constellation, creativity, IQ, and academic achievement. *Child Development, 38*, 481-490.

Connor, J.M., & Serbin, L.A. (1985). Visual-spatial skill: Is it important for mathematics? Can it be taught? In S.F. Chipman, L. R. Brush, & D. M. Wilson (Eds.), *Women and mathematics: Balancing the equation* (pp. 151-174). Hillsdale, NJ: Lawrence Erlbaum.

Dinnel, D.L., & Glover, J.A. (1987, April). *Processing differences between competent and less competent mathematics students*. Paper presented at the American Educational Research Association, Washington, DC.

Dossey, J.A., Mullis, I.V.S., Lindquist, M.M., & Chambers, D.L. (1988). *The mathematics report card: Are we measuring up? Trends and achievement based on the 1986 National Assessment*. Princeton, NJ: Educational Testing Service.

Eccles, J.S., & Jacobs, J.E. (1986). Social forces shape math attitudes and performance. *Signs: Journal of Women in Culture and Society, 1*, 367-380.

Educational Testing Service. (1971). *Sequential tests of educational progress (STEP, Series II): Teacher's manual for administering, scoring, and interpreting grades 10-12*. Princeton, NJ: Author.

Fennema, E., & Peterson, P. (1985). Autonomous learning behavior: A possible explanation of gender-related differences in mathematics. In L.C. Wilkinson & C.B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 17-35). Orlando, FL: Academic Press.

Fennema, E., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization, and sociocultural factors. *American Educational Research Journal, 14*, 51-71.

Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in mathematical tasks. *Review of Educational Research, 59*(2),185-213.

Green, D.G. (1987, August). *Sex differences in item performance on a standardized achievement test battery*. Paper presented at the meeting of the American Psychological Association, New York. (ERIC Document Reproduction Service No. ED 291806)

Hawn, H.C., Ellet, C.D., & Des Jardines, L. (1981, April). *Differences in mathematics achievement between males and females in grades 1-3*. Paper presented at the annual meeting of the Eastern Educational Research Association, Philadelphia.

Hedges, L.V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107-128.

Hedges, L.V., & Olkin, R. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hilton, T.L., & Berglund, G.W. (1974). Sex differences in mathematics achievement: A longitudinal study. *Journal of Educational Research, 67*, 232-237.

Huston, A.C., & Carpenter, C.J. (1985). Gender differences in preschool classrooms: The effects of sex-typed activity choices. In L.C. Wilkinson & C.B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 143-165). Orlando, FL: Academic Press.

Hyde, J.S. (1981). How large are cognitive gender differences?: A meta-analysis using w^2 and d. *American Psychologist, 36*, 892-901.

Hyde, J.S., Fennema, E., & Lamon, S.J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139-155.

Jacklin, C.N. (1989). Female and male: Issues of gender. *American Psychologist, 44*, 127-133.

Jones, L.V., Burton, N.W., & Davenport, E.C., Jr. (1984). Monitoring the mathematics achievement of black students. *Journal for Research in Mathematics Education, 15*, 154-164.

Kaiser, H.F. (1958). The Varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23*, 187-200.

Kimball, M.M. (1989). A new perspective on women's math achievement. *Psychological Bulletin, 105*, 198-214.

King, E.M. (1977). *Canadian tests of basic skills, metric edition, levels 9-14, form 4M, teachers guide*. Toronto, ON: Nelson.

King, E.M. (1982). *Canadian tests of basic skills, high school, multilevel edition, levels 15-18, form 5, teachers' guide*. Toronto, ON: Nelson.

Lewis, J.C., & Hoover, H.D. (1986, April). *Differential prediction of academic achievement in elementary and junior high school by sex*. Paper presented at the annual meeting of the American Educational Research Association, Montreal.

Linn, M.C., & Hyde, J.S. (1989). Gender, mathematics, and science. *Educational Researcher, 18*(8), 17-19, 22-27.

Lips, H.M. (1988). *Sex and gender: An introduction*. Mountain View, CA: Mayfield Publishing.

McKay, L.R. (1979). Sex differences in cognitive styles and mathematics achievement in fourth and eighth graders. (Doctoral dissertation, University of Texas at Austin, 1978). *Dissertation Abstracts International, 39*, 4103A.

Myerberg, N.J. (1986). *The annual report of California Achievement Tests results, 1985-1986*. Rockville, MD: Department of Educational Accountability, Montgomery County Public Schools. (ERIC Document Reproduction Service No. ED 286894)

Purkey, W.W., Cage, B.N., & Graves, W. (1973). The Florida key: A scale to infer learner self-concept. *Educational and Psychological Measurement, 33*, 979-984.

Ramist, L., & Arbeiter, S. (1986). *Profiles of college-bound seniors: 1985*. New York: College Entrance Examination Board.

Randhawa, B.S. (1988). Basic skills and macro- and micro-analysis of mathematics achievement of grade ten students as a function of gender and locale. *British Educational Research Journal, 14*(2), 141-148.

Randhawa, B.S., & Beamer, J.E. (in progress). Sex differences in first year university mathematics.

Randhawa, B.S., Beamer, J.E., & Ng, C.F. (in progress). Sex differences in mathematics in relation to contextual variables.

Randhawa, B.S., & Hunt, D. (1987). Sex and rural-urban differences in standardized achievement scores and mathematics subskills. *Canadian Journal of Education, 12*, 137-151.

Raphael, D., Wahlstrom, M., & McLean, L. (1984). Results from the second international mathematics study: Are boys better at math than girls? *Orbit, 15*(2), 6-9.

Schonemann, P.A. (1966). A generalized solution of the orthogonal procrustes problem. *Psychometrika, 31*, 1-10.

Swafford, J.O. (1980). Sex differences in first-year algebra. *Journal for Research In Mathematics Education, 11*, 335-346.

Tsai, S.L., & Walberg, H.J. (1979). Mathematics achievement and attitude productivity in junior high school. *Journal of Educational Research, 76*, 267-272.

Weiner, N.C. (1984). Cognitive aptitudes, personality variables, and gender difference effects on mathematical achievement for mathematically gifted students (Doctoral dissertation, Arizona State University, 1983). *Dissertation Abstracts International, 44, 3621A.*

Wilder, G.Z., & Powell, K. (1989). *Sex differences in test performance*. (College Board Report No. 89-3). New York: College Entrance Examination Board. (ERIC Document Reproduction Service No. ED 308331)

Wittig, M.A., Petersen, A.C. (Eds.). (1979). *Sex related differences in cognitive functioning: Develoemental issues*. New York: Academic Press.

INTERCHANGE

A QUARTERLY REVIEW OF EDUCATION

- a leading international journal on the critical issues facing educators in Canada and abroad
- a recipient of the prestigious CAFE award for the most original work in educational foundations
- four issues per year feature scholarly articles, critical notices, book reviews, and debates
- recent special issues include Creativity, Thought, and Mathematical Proof; History, Science and Science Teaching; and Understanding Literacy (guest edited by David Olson)
- Edited by Ian Winchester

Regular One-Year Rates for 1990: Institutions $50.00; Individuals $35.00. For 1991: Institutions $55.00; Individuals $35.00. To order send cheque, purchase order, or complete MasterCard or VISA information to OISE Publishing, 252 Bloor St. W., Toronto, Ontario, Canada M5S 1V5.



Published by The Ontario Institute for Studies in Education, Toronto, Canada

The Alberta Journal of Educational Research *Vol. XXXVII, No. 3, September 1991, 259-275*

PERSPECTIVES

The Burt Controversy: An Essay Review of Hearnshaw's and Joynson's Biographies of Sir Cyril Burt[1]

JOHN HATTIE
University of Western Australia

When the controversy relating to Sir Cyril Burt broke in the mid 1970s, the academic climate was most conducive to a scandal. Sir Cyril Burt was associated with defenders of the English class system; many were criticizing those who believed heredity was a major contributor to individual differences; the furor had not abated regarding the 1969 Jensen article which included comments about the genetic reasons for black-white IQ discrepancies and had used Burt's work; and the issues of race, genetics, and education were mixed with politics and demonstrations. Many academics were accused of racism and elitism and of concocting data to support the status quo; that is, the hereditarian viewpoint. It is perhaps not surprising that someone who was would be accused of fraud, especially if that someone was disposed toward the heredity view, seen as elitist, of the establishment and traditional school, and preferably dead.

The incidence of fraud is not known, although over 90% of academics in one study cited instances of cheating known to them (Saint John-Roberts, 1976). Broad and Wade (1982) claimed it was far less, but not negligible. Among them Broad and Wade (1982), Gould (1981), Grafton (1990) and Kohn (1986) cite over 100 cases of academic fraud. No matter what the incidence, it takes only a few cases to be publicized for the credibility of science to be damaged. Burt became psychology's biggest test.

The Accusations

The rumors about Sir Cyril Burt, Britain's most famous psychologist, began in the early 1970s. Jensen (1974) published an article questioning some of Burt's data on twins, but the rumors were fed mostly by Kamin in his various talks and subsequently by his book (1974). Kamin's book overstated the case against the heredity viewpoint and included virulent attacks on those who supported it, particularly Burt, although he stopped short of accusing Burt of fraud. He concluded that "the numbers left behind by Professor Burt are simply not worthy of our scientific attention" (p. 47). When a London *Sunday Times* journalist broke the story on the front page on October 24, 1976 with

John Hattie is a professor of education. His research interests are in cognition, creativity, psychometrics, and meta-analysis.

the use of the terms "fraud," "false data," and "invented crucial data," the academic community could no longer stay silent.

There were many accusations and counteraccusations that immediately followed, but it is notable that most of these lacked evidence. It was known that Hearnshaw,* an established British historian of education who had begun to write an official biography of Burt, would now include a commentary on the "scandal." Hearnshaw had the necessary competencies to write the biography—a capacity for absorbing facts, a capacity for stating them, and a point of view (Strachey, 1931).

He made six major claims about Burt. First, Burt could not have possibly collected some of his postwar data and thus he fabricated them by inventing data to fit his theories. Hearnshaw (1979) cited McAskie (1978) who claimed that Burt's data was suspiciously close to normal distributions and argued that Burt's "besetting weakness was to rely on statistical manipulation rather than empirical investigation, forgetting that poor data cannot yield sound conclusions" (p. 70). Second, Burt's crude estimates of parental IQ in the 1920s were passed off as more accurate in the 1960s. The suggestion was that the estimates were adjusted to accord with Burt's preconceptions. Hearnshaw concluded that "the table of IQ scores and social class gradings was an elaborately constructed piece of work, and we are forced to the conclusion that he simply did not possess detailed data, at any rate for the whole sample of his separated MZ twins" (p. 247). Third, correlations changed when the sample sizes did not, and correlations remained invariant (to the third decimal point) when the sample sizes changed. Hearnshaw concluded that the accounts "given in Burt's published papers were false; and that a measure of deception was certainly involved" (p. 241). Fourth, some of Burt's co-workers were fictitious. The claim is that some of his co-workers (particularly Howard and Conway) were apparently invented to justify his collecting data, or so that Burt could further promulgate his views, defend some of his priority claims and boost the stature of his work. Fifth, Burt changed other researcher's contributions to the journal he edited to reflect better his views and contributions. Evidence is provided that Burt did this with the contributions of Eysenck, Clarke and Clarke, and Wrigley. Sixth, Burt altered the history of factor analysis to claim priority over some contributors. Hearnshaw documents the case that Burt arrogated to himself the whole of Spearman's fame, became increasingly obsessed with questions of priority, and unquestionably misrepresented Spearman's position. Moreover, Burt gave Pearson pride of place such that "Spearman was dethroned, and Burt himself elevated in his place! It seems almost a classic example of the killing of the father king" (pp. 176-177). "He falsified history in the interest of self-aggrandizement. That he was guilty of malfeasance there can be no reasonable doubt. The only question at issue can be, was he guilty *simpliciter* or guilty with diminished responsibility as a result of pathological influences?" (p. 180). Hearnshaw answered the question by supporting the latter argument.

*Unless otherwise stated, references to Hearnshaw and Joynson refer to Hearnshaw's 1979 book and Joynson's 1989 book.

Hearnshaw concluded that the data Burt used for his calculations were poor and unreliable, that he made a great many unexplained, careless and inconsistent adjustments to the raw scores of his tests with the results that the figures are quite improbable. He ultimately applied sophisticated statistical techniques to scientifically almost worthless data, with disastrous results. In three instances, Hearnshaw finds Burt guilty of deception: He falsified the early history of factor analysis; he produced spurious data on MZ twins; and he fabricated figures on declining levels of scholastic achievement.

Burt's Motives

Hearnshaw traced Burt's pathology from the early "marks of a delusional system, growing from small beginnings into a blind and warping compulsion" (p. 177). Hearnshaw attributed this pathology to seven causes.

First, Burt's onset of Meniere's disease, diagnosed in 1941, tended to incapacitate him. Hearnshaw noted that those with Meniere's disease typically were obsessional compulsive types with marked psychosomatic lability. Second, Burt's competencies were declining with age. Burt complained of "senile paragraphia" (surely this was a joke!) and there are isolated instances of misplaced words in his notes and letters to his sister. As he was often his own editor, Hearnshaw claimed that he was deprived of a very necessary corrective. "Carelessness, then, partly the result of haste and partly the result of emotional involvement and declining powers, was assuredly a contributory ingredient in the final product" (p. 242). Third, there were warnings in his early childhood as he was forced to survive by observing his peers and enemies closely, "to keep his feelings to himself, to bluff it out, and to outmaneuver those who tried to molest him" (p. 273). "In the lost boyhood of Cyril Burt psychology was betrayed" (p. 291). Fourth, he was suffering other setbacks: his marriage, the breakdown of his health, his difficulties following retirement from the chair, loss of control over the statistical journal he helped found and nurture; the gradual erosion of the system of selective secondary education that he had done so much to promote; new modes of thinking and younger rivals were ousting him from center stage; and the doctrines he believed in were being rejected. These setbacks, according to Hearnshaw resulted in his gamin personality (living by treachery and underhand trickery typical of London slum children) coming to the fore and he could appear "quite unscrupulous in the expedients he employed" (p. 276). Fifth, many of Burt's papers were left at University College during the war, and subsequently were bombed in the blitz. Burt subsequently invented the "lost" data. Hearnshaw is certain that all Burt's data were lost (p. 250) but latter he claimed that the greater part was not destroyed (Hearnshaw, 1990). Sixth, Hearnshaw claimed that Burt collected no new data after 1955 as he had neither the physical fitness, the assistants, nor the resources to carry out more research. Data attributed to post-1955 were fabricated. Seventh, Hearnshaw charged that Burt could be slippery and unscrupulous and had an egomaniac demeanor. He would "misrepresent opponents and blur the issues. Beneath a polite exterior and apparent reasonableness was a steely determination to get the better of the argument, and to humiliate his op-

ponents" (p. 206). Moreover, he could be "hostile, cantankerous, and, if need be, unscrupulous" (p. 270).

Overall, Hearnshaw's biography appeared to be a letdown for the great man: a letdown from preeminence to scoundrel, from innovator to the delinquent. Hearnshaw convinced the majority and the reviews were full of praise. It was, claimed Broad and Wade (1982), a sympathetic and subtly drawn portrait, and Gould (1981) called it splendid.

The Counterclaims

Hearnshaw's views were unchallenged for a decade, until Joynson (1989) published his book which concluded that "the evidence which has so far been presented is insufficient to support the accusations which have been made. It does not follow that Burt's behavior was invariably commendable, and that he possessed no weaknesses of character. But the gross misbehavior of which he has been so widely accused has not, in my opinion, been established. A grave injustice has been done" (p. ix).

Joynson was particularly critical of Hearnshaw, as unlike Flaubert's ideal biographer who is "everywhere present and nowhere present," Hearnshaw does not stay in the background. Joynson claimed that Hearnshaw used flimsy evidence, did not locate primary references, has a tenuous grasp of Burt's procedure, omitted words in quotations that lead to a distorted picture, depended on gossip about long past incidents entirely unsupported by adequate evidence, does not know what Burt is talking about, failed to treat evidence critically, used unverifiable speculation about superfluous absurdities, too often accepted hostile comments of Burt's enemies without examination, was selective of evidence, was one-sided, and made innumerable errors. Joynson offered countering evidence for each of the six major charges made by Hearnshaw.

Regarding Burt's guessing of parental and occupational IQ and then passing these off as real, Joynson agreed that Burt often gave less information than was desirable, but contrary to Hearnshaw's claims, Burt did acknowledge the unreliability of his data with comments such as "pilot enquiry," "less thorough and less reliable" and pointed out to correspondents the problems of these data. It seems to me that Burt on more than one occasion noted the unreliability of these data, using terms such as the "comparisons are highly precarious," the figures cited are "the best we can get in the way of objective data," "this caution (the difficulty of procuring samples which can be safely regarded as comparable) applied equally to the figures set out above," and we need to "make liberal allowance for" the way the data were collected. Researchers still estimate parental IQ in ways similar to Burt's procedures (see Jensen, 1974).

Joynson did see some case for criticism of Burt's (1969) description and presentation of material in his "Intelligence and Heredity" article. He explained, however, how Burt could have obtained his data and noted that the figures that Hearnshaw claimed were fabricated are precisely those that have the least tendency to support Burt's argument. "If Burt had been inventing the figures he would surely have chosen numbers which would strengthen his argument rather than weaken it" (p. 209). It seems that Burt

collated means from various publications by himself, his research students (many of whom he acknowledged were now teachers), and perusing manuals of tests, such as Schonell, Vemon, Weschler, or the National Foundation for Educational Research. This crude attempt at meta-analysis shocks Hearnshaw but was a precursor to a revolution of the 1980s (Glass, McGaw & Smith, 1981). Many footnotes and references, not discussed by Hearnshaw or Joynson, outlined where Burt could have gleaned the data. For example, Hearnshaw considered that Burt could not have obtained the 1955 data as "by then Burt had no right of entry to London schools" (p. 257) and concluded that these data were fabricated. Burt detailed (perhaps overdetailed) the collection of such data in 1954 with the help of Mr A.M. Reid, of over 300 London boys at the age of 9.0-10.0, and four years later retested the same group. Further, Audley and Rawles (1990) have claimed that Burt had access to data gathered by the National Foundation for Educational Research from 1955 to 1965 for a survey of ability and attainment. It is claims like these that lead to difficulty in presuming that Burt fabricated his data.

That the correlations suspiciously changed while the sample size did not, or remained invariant when the sample sizes did change, appears to be the hardest claim to overturn, and I confess it took me several readings of this chapter to understand fully Joynson's explanations. He first corrects some anomalies such as typographical errors noted by Burt (Ron Hambleton kindly lent me his copies of Burt's papers which he obtained in response to an advertisement from Burt's secretary shortly after Burt's death. These offprints include corrections and support Joynson's claims. Also see Jensen, 1974, Skanes, 1978). Joynson then argued that the .77 correlation for MZ twins reared apart in 1943 and the .771 in 1955 may be different although the sample size increases from 15 to 21, as the former could be anywhere between .766 and .774; that the .54 for correlation for DZ twins in 1943 and .542 in 1955, though the sample size increased from 157 to 172, could be the result of cumulative addition of new data to the previous data. Thus both are plausible. This is difficult to believe, although not impossible. There is a more obvious conclusion, though it too presents a problem. Burt (1955) cited 21 *pairs*, but Conway (1958) cited 42 *cases*, so the data are probably the same. The problem is that the correlations changed (although the mean change is only .008). In Burt (1966a) it is clear that there are 53 *pairs* (mean change compared to 1955 is .016). It seems believable that Burt could have located an extra 32 pairs in the 11 years (particularly if his acknowledgments to students and colleagues are to be believed, e.g., Burt, 1966a, p. 141).

All anomalies are traced to two papers of which 29 correlations are invariant out of the 64 reported. The invariant correlations, however, "are not to be found concentrated in the crucial areas of tests of intelligence, nor in relation to MZ twins, as might have been expected if Burt was primarily concerned with fabricating evidence about the inheritance of intelligence of twins. Rather, they appear predominantly among the physical measurements, and are scattered across every category of subject" (Joynson, p. 146). Only four concern intelligence, of which two have already been questioned, above. Moreover the coefficients changed in the direction opposite to Burt's hypothesis: "the changes rather suggest a strengthening of the en-

vironmental influence, a strange effect for a scheming hereditarian to produce" (p. 159).

If Burt wanted to stretch his case he would have changed the intelligence coefficients; he would not have been so stupid as to leave such obvious consistencies, would have faked more carefully, and would have changed the correlations in the direction of his hypotheses: "A guilty man would never have made such an error" (Joynson, p. 157). If he had faked them, why did he not point to the remarkable consistencies as evidence in his favor? Why invent numbers and then not reap the rewards? Thus Joynson concluded that there has been no reliable indication of fabrication whatever.

With respect to the existence of Burt's co-workers, Joynson cited Burt's housekeeper that Burt was fulfilling a promise to credit the "missing" co-workers with work done earlier, long before he wrote the articles. Joynson agreed with the critics that Burt frequently wrote letters, comments, and book reviews and published them in his journal under various pseudonyms. He dismissed the claim, however, as an unprofitable basis for sober conclusions but an admirable subject for sensational speculation in the media. It may be possible to attribute some of the sole authored articles by Howard and Conway to Burt (e.g., by computer concordance procedures), although it is probably not worth the trouble. After reading the various contributions by Howard and Conway, what is striking is not the bias and support of Burt's contributions as claimed by Hearnshaw, but the opposite. The papers and reviews are sober, well argued, extensively footnoted; and the style is unmistakably extremely close to Burt's.

Howard has been sighted by some of Burt's peers, by at least one twin who was tested by her, and according to Burt she emigrated before 1950. It is possible that she is the same Howard who was once a member of the British Psychological Society. That no one answered a request in the London *Times* is far from convincing evidence that she or Conway do not exist. It is easy to imagine many people in the back blocks of Britain, Canada, or Australia who have yet to receive their 1976 editions of the *Sunday Times*; and it is possible that Miss Howard and Miss Conway married and changed their last names. As Hearnshaw claimed, they were probably LCC social workers or care committee workers and thus it is not surprising that they would be difficult to locate.

That Burt changed others' contributions to the journal he edited is not challenged; whether these changes aimed to reflect better his view is. It is probable that Burt performed his editorial duties with relish, but most of the documented cases are trivial, misleading, or false. For example, one case documented by Hearnshaw is that of Clarke, who claimed that Burt had altered his and his spouse's work such that they "did not recognize them as the same summaries ... [and] the whole emphasis of 'our' articles was slanted (against Eysenck)" (p. 148). Joynson countered that when compared with the actual theses, the published summaries (not articles) "appear to me to be accurate and well-balanced accounts of the main points contained in the theses" (p. 246), and if changed at all it was changed to be more favorably inclined toward Eysenck than was the thesis. Changing others' work or claiming credit for others' work is a common claim, especially of students

about mentors. This is an often discussed issue that is not unique to Burt. There were also many who were appreciative of the attentions that Burt gave to their submissions. Valentine (1965) paid "tribute to Burt's generosity in helping others engaged in psychological research.... he has put many a young psychologist on his first path to advancement ... [his] elaborate discussions, which were most useful, but embarrassed me with the thought of the time he must have spent" (p. 20).

Joynson was most critical of Hearnshaw's claim about Burt altering the history of factor analysis to claim priority over some other contributors. Joynson noted that Burt differed from Spearman on some major issues; and could not ever be accused of disagreeing with Spearman only after the latter died. Burt, like many others, was not as convinced that the existence of a general factor excluded group factors, and Burt was far more interested in specialized abilities than Spearman. Joynson can find no reference to Pearson (1901) anywhere in the papers where Burt supposedly usurped the origins from Spearman. He argued that Hearnshaw failed to appreciate the enmity between Spearman and Pearson which Burt probably did not wish to enter, and that Burt was more influenced by Macdonell (1901) than Pearson. He contended that the reasons for Burt's renewed claims about the origins of factor analysis were probably related more to the publication of Thurstone's seminal work (1947) rather than to Spearman's death in 1945. Thurstone's book tended, in Burt's view, to ignore the contributions of Burt and other British factor analysts. This is indeed the case (compare Cronbach 1979) with respect to the importance of group factors as well as a general factor, and the priority of the invention of the centroid method. I do agree with Thurstone, however, that Burt's main omission (and possibly his lack of later influence) related to his lack of desire or belief in rotation as an aid to the interpretation of factors. Joynson clearly demonstrated that these claims of priorities can be traced to Burt's published work as early as 1909. I also note the lengthy footnote 3 in Burt (1949) which explained why so many researchers are "naturally unfamiliar with these early developments," why Pearson should get credit, and reference to a debate in 1910 between these various academics about priority.

There are still some who would contend that Pearson made the initial major contribution to the development of factor analysis, and there are peculiar histories of factor analysis that make Burt's beliefs seem mild (e.g., Schonemann, 1987; Steiger & Schonemann, 1976). Thurstone (1947) made sweeping claims about the British school (and generally omitted their influence), and Guilford (1936) omitted Burt altogether. Much of the historical debate can be traced to the differences between factor analysis (initiated by Spearman) and derivatives such as centroid and principal component analysis (initiated by Pearson). Controversy between whether there are, and the nature of, differences between these procedures still rages (Velicher & Jackson, 1990). Further, it needs to be noted that Burt was "foremost a practical applied psychologist," as Hearnshaw acknowledged (pp. 35, 92) and in this light many of Burt's claims are not altogether incorrect. Burt was the first, or among the first, to *use* some factor methods and this first *use* of factor methods is the consistent theme in his writings (e.g., Burt, 1968, p. 67).

Joynson concluded that there are so far no grounds for accusing Burt of historical falsification or lying; the critics have not established their case conclusively in a single instance, and "any delinquencies of which Burt might be at all plausibly accused appear to have been of a minor and relatively common variety, and come nowhere near the sensational frauds which have been so widely alleged" (p. 214). Hearnshaw claimed that the motive for the usurpation was self-aggrandizement, whereas Joynson suggested that it was more likely ambition or self-assertion, and when Burt's motives are labeled self-aggrandizement, then the accusers are already halfway to accepting his guilt. Burt, claimed Joynson, ascribed first place to another, was the most dependable supporter of Spearman's central idea (I do not agree with this), was pushing for British priority as well as his own, and could have pushed his claims more forcefully rather than in his modest and polite ways. Joynson questions why this belittling took 40 years to note and answers that it was because "there was no such campaign. Burt's views were put forward quietly, and unsensationally.... The notion of a mounting campaign against Spearman is a fiction of Hearnshaw's imagination. Burt was a sane man pursuing a rational purpose" (p. 125). This is not to contend that Burt was unbiased, as history is open to numerous shades of interpretation. The latter account of his own earlier papers appears to be substantially accurate, as he was never entirely derivative from Spearman. Hearnshaw has "painted the contrast too starkly, greatly exaggerating both the extent to which Burt was originally dependent on Spearman, and the nature and strength of his more independent behavior in maturity" (p. 109).

The Merits of the Cases

Joynson is convincing when he claims that the majority of Hearnshaw's arguments employ negative evidence: If Burt had possessed the data, he would have replied, but he did not reply so he probably did not possess them; moreover, one would have expected there to be traces of assistants, records of data, records in diaries, but such evidence is not found. Joynson's reaction to Hearnshaw's claims as to causes of Burt's indiscretions are convincing. It is difficult to support the notion that because Burt had Meniere's disease his critical judgment became impaired; he may have been more frustrated and thus may have become more careful. Joynson finds it remarkable, as do I, that Hearnshaw could claim that "there was in all probability an innate instability in Burt's psychosomatic make-up" (p. 277). A further claim of Hearnshaw borders on the ludicrous: "by reason of its exclusiveness (classical education) tends to induce feelings of superiority and elitism in those who are classically educated; and in a subtle way it is a training in pretence" because in preparation of producing pieces of Greek and Latin prose the perfect copy or fake achieves the highest marks. "For more than ten years of his early life Burt was subjected to this regular discipline" (pp. 263-264). While some of these claims by Hearnshaw may account for part of Burt's personality, it is beyond the scope of even a psychobiography to believe that this could explain why he allegedly cheated and so many others did not. Joynson's main contention was that Hearnshaw "makes all those assumptions which are consistent with the interpretation he wants to make, conveniently ignoring any pos-

sibility which might conflict" (p. 221). Hearnshaw selected only setbacks and does not account for the successes (such as honorary degrees, knighthoods, success of articles received, laudatory comments by colleagues).

Joynson is not believable in inventing excuses for Burt's failure to reply to requests for data (p. 175), claiming that the correlations for the measures of intelligence of the twins were not markedly stable when the mean difference appears to be very stable (mean difference = .043, sd = .038), that the misleading claims in a newspaper article may have been a consequence of Burt having misheard the question, that his critics are part of a left-wing, even Communist, conspiracy or part of a Maudsley cabal, and that the subject is moot because one cannot convict Burt of fabrication on suspicion, especially since he cannot defend himself.

Joynson is on the weakest ground when he tries to psychoanalyze Hearnshaw. Joynson is at pains to understand how Hearnshaw so misread Burt's character. Hearnshaw had already been working on the biography for some five years, and had written several chapters. Then along came the critics and Hearnshaw had to revise many opinions. This, claimed Joynson, added considerable pressure on Hearnshaw to complete his inquiries as soon as possible. Given the pressure, he hit upon the idea of striking a compromise between the anti- and pro-Burters, which "inevitably meant a verdict of guilty—not as guilty as the anti-Burters alleged, but not as innocent as the pro-Burters maintained" (p. 322). Joynson patronizingly commented that "a more leisurely timetable would have been wise" (p. 306). Joynson's attack on Hearnshaw is the worst feature of an otherwise excellent rebuttal.

The Reaction

There have been reviews of Joynson's book and the overriding opinion is bouquets for the detail and the way Joynson has corrected the little matters, accolades for having the bravery and time to try, and muted applause for edging the opinion slightly more to the side of innocence. The key players, however, are not convinced that Joynson has moved the opinion sufficiently to cross the line of innocence, and some have claimed that he is not even close.

Clarke and Clarke (1990a) are not convinced. They are not happy with Joynson's methods as he "unfortunately uses every available means, including gossip and innuendo, to cast doubt upon the integrity and scholarship of Burt's critics" and "in his zeal to vindicate Burt, Joynson has been less than scrupulous in his search for, and assessment of, the considerable documentary evidence," and he has "made many errors of fact" (p. 123). As in their earlier accusations, they provide no evidence for any of their claims. They now acknowledge that the changes supposedly made by Burt to their own work were "subtle" which greatly diminishes their case (Clarke & Clarke 1990b). Their reviews can be dismissed as shallow attempts to cast doubt on the integrity and scholarship of Joynson.

Beloff (1990), the organizer of the British Council symposium on Burt that aimed to "decently face the evidence before it," also is not convinced and contends that "it is a pity that [Joynson] does not take a fresh look at the problem, considered from a different social and ideological perspective." It is

difficult to imagine two more unlikely perspectives than Hearnshaw's and Joynson's. Beloff casts aspersions on Joynson's scholarship with condescending comments such as "it must be rare indeed for a contemporary work [i.e., Hearnshaw's] to receive such detailed textual analysis," "Joynson's emotive argument, induces despair more often than clarity," and she concluded that Burt "could not, nor can now, obscure the doubt, beyond reasonable doubt, that the later evidence was not empirical" (she means that Burt invented data). Her review is a poor apology for the 1970-1980s establishment of British psychology.

Two key players have modified their views. Eysenck (1990) who had already changed his views from innocent to guilty, now concludes that the most acceptable verdict is not guilty. He regards the case of the "missing ladies" as comic opera stuff and, correctly, does not take it seriously; he agrees with Joynson on the legitimacy of Burt's claims to priorities in factor analysis; but he believes that Burt did invent at least some of his twin data. Hudson (1989), who was responsible for the decision to burn Burt's archives and was an early accuser of Burt, was most concerned that Burt's transgressions were patently obvious but not discussed at the time, and he now seems to soften his view as to Burt's guilt.

Zenderland (1990) appears persuaded by Joynson's claims but concludes that Burt hardly leaves the courtroom with his reputation intact. She agrees with Joynson that Burt's methods may have been less than admirable, but they were short of criminal. She is not convinced by some of Joynson's claims about the kinship correlations, and believes that Joynson's "innocent" Burt emerges as an even less likable character than Hearnshaw's "guilty" Burt. Fortunately, this is not a popularity contest. Blinkhorn (1989) agreed with the claims about Burt's priorities regarding factor analysis and his elevation of Pearson, but was not convinced about Joynson's claims about the twin data. He found merit in the argument that it is harsh to judge pioneers by standards they ultimately set after their fieldwork was completed, and wonders if others are going to leave self-documented archives for future defenders. Fancher (1989) does not agree with the latter and sees less honor in Burt's deviousness. Some reviewers merely summarized the book (e.g., Ammons, 1990; Stein, 1989), approved the reinvestigation of the matter (Thieman, 1990; Tucker, 1990), or agreed that the question marks that Joynson has created deserve some response (Audley & Rawles, 1990; Plomin, 1989; Tysoe, 1990).

Hearnshaw (1990) and Joynson (1990) have also clashed since the publication of the latter's book. Hearnshaw accused Joynson of introducing no new evidence, not examining available evidence, not consulting him, not knowing Burt personally, and Joynson's book is of tedious length (Joynson is 27 pages shorter). These claims are either incorrect or not damning. Hearnshaw concluded that "Joynson's attempt to salvage Burt's reputation may be regarded as a total and dismal failure, and it fails above all because Joynson has simply not done his homework" (p. 64). Joynson (1990) reacted to these claims, but the details add little to the overall story. The editors of the British Psychological Society official magazine ruled in October 1990 that because they had already devoted considerable space to the Burt affair they

would not publish Joynson's rejoinder to Clarke and Clarke (1990b). They continued, however, to publish other letters on this subject in subsequent issues. They have also devoted much space to discussing Joynson's claim that Burt was the only psychologist to be knighted (the number now stands at seven plus a dame).

The Verdict

Both books are necessary, compelling reading. Both are extremely well organized and a sense of the author and subject are vitally present. Hearnshaw's biography has more flesh and bones about Burt and there is an overall persuasive argument as to Burt's guilt of fraud. At times Joynson protests too much and his attacks on the motives of the accusers detract from his overall message. Joynson has introduced the issue of whether Burt was guilty "beyond reasonable doubt," but he does not, nor do any reviewers, raise the issue as to whether only a taint or a full proof is needed for fellow researchers to discard Burt's data. Neither book is convincing such that a closed case exists. Like most historical mysteries, this is maybe how it should be.

So whom are we to believe? Joynson claimed that many believed Hearnshaw, because he enjoyed a reputation as a reliable and balanced historian. Further, he had reached a definite conclusion that satisfied most with the verdict of guilty as accused, as this "offered the best hope of a quick settlement and a speedy oblivion" but did not find Burt as guilty as some had been commenting or as innocent as others believed. The reviewers of Hearnshaw's book were laudatory, agreed with his conclusions (many were written by Burt's initial accusers, which Joynson claimed was reasonable except "in so far as their ill-concealed pleasure at the outcome outweighed their capacity for critical appraisal," p. 314), were brief, did not show much critical insight, or did not make much penetrating analysis of the book's strengths and weaknesses. Other reasons for support of Hearnshaw's biography included: Burt's most influential supporters changed their minds when the biography appeared; further papers, published after Hearnshaw, claimed more evidence against Burt; and the Council of the British Psychological Society condemned Burt on the basis of the biography. That the Council rushed into assuming guilt, particularly when two of the main accusers (Tizard and Clarke) were major office-bearers, still seems rather unseemly.

The impression of Burt often left by Hearnshaw and Joynson (and certainly reflected in the reviews and second-hand accounts) is that of slipshod practices, lack of detail, and a mind closed to and oblivious of opposite arguments. In preparation for this essay I have read most of Burt's work and I was struck by the detail, the lucidity, and the competence that emanates from it. We are dissecting the works of one of the clearest writers in psychology, and certainly one of the most literate and artful. Although Burt may have had an egomaniac demeanor, he is not alone in academia in this quality. Although I never met Burt, it seems reasonable to infer that his commanding manner was not endearing and his attacks not welcomed. During the past year I have read biographies of McLuhan, Blunt, Freud, and Mead and note many similarities in personality, and could name many living exemplars with

similar egomaniacal demeanors to Burt. Moreover, what would we make of an eminent psychometrician who promises to call his autobiography "Mean-square within"? Burt may not have been a sociable or clubable man (Hearnshaw, p. 183), but surely these indicators are not symptomatic of a criminal mind. It seems to be a serious charge to be accused of not being part of the club in Britain.

So was Burt guilty of anything? Burt could have collected some of his postwar data with the help of ex-students; he was honest about the low reliability of his data, and his caveats were readily available to those not convinced or willing to read his references; he probably received assistance from Conway and Howard although he used their names as by-lines, probably without their directly assisting him at the time of writing; he almost certainly changed contributors' submissions to the journal and bullied some ex-students and colleagues; he does deserve more credit for his pioneering use and developments in factor analysis; and he presented an argued (although arguable) case for the priority of Pearson's works. The case of the robust correlations is more difficult to reconcile, although it appears that it was a case of carelessness more than deceit or intention to commit fraud. To prove fraud, there must be a case for intention to deceive, and this is Hearnshaw's weakest point.

It does appear that Burt committed awkward, unskilled acts, but there is little evidence that he committed intentional deceit or fraud. Burt disappointed many, particularly with his devious replies, and at a minimum he can be accused of occasional slipshod work and that he took inappropriate liberties; at worst, we can ignore his work. To go further awaits more evidence.

The Moral

A number of lessons can be learned from this unfortunate affair. The need to be pure is the essence and there is no excuse for inventing data, changing conclusions (in any direction), or plagiarizing. It is not good enough to comment that comparatively few "can resist the temptation to indulge in a little faking, fishing, and even bare-faced prevarication, when things go wrong. But are not all of us tempted at times in much the same way? ... even scientists themselves often display different standards of truthfulness in different situations" (Burt, 1966b, p. 367). Students need many reminders of the importance of not comitting these sins. Starting as undergraduates they need to be reminded about the evils of copying from texts without reference, passing off another's argument as their own, not acknowledging sources (even if the source be their peers), trimming data to make it look more convincing, embellishing results, omitting aberrant values, misreporting methods, selecting only the "best" results, or hiding failed experiments. As I have traveled around universities during my present sabbatical leave, I have asked to see the guidelines and penalties for these offenses. The most notable feature is the lack of guidelines, their inaccessibility *prior* to committing any of these acts, and the minimal nature of the penalties. I am also aware of the difficulties of proving many of these issues to a panel of independent academics (and usually to fellow student juries) and thus see a stronger import for the clear—and available—delineation of the rules and procedures

for investigating fraud. It should also not be underestimated how difficult it is to prove fraud or innocence of it and/or to convince university authorities to investigate (Broad & Wade, 1982, document numerous cases where authorities were reluctant to investigate).

There is no doubt that Joynson is correct in asking for procedures to deal with false accusations. When one is wrongfully accused of fraud (or racism or sexism), it is extremely difficult to mount evidence against the claims, as most efforts lead to supporting the very claims. Most do not generate positive evidence and the onus is on the accused to prove innocence. In future, it is hoped that allegations are published along with the evidence and such claims open to serious academic debate, preferably prior to being popularized in newspapers. Any contentious area needs protection so that the advocates may say their piece, and there is a need for outlets for critical appraisal of these views. More published interchanges (such as exist in *Brain and Behavior* and *Interchange*) are needed.

More attention needs to be placed on popularizers of claims of fraud. The need for simplicity and everyday language can lead to misrepresentation, although it is not reasonable, as does Joynson, to criticize newspaper reporters for wanting to get a "big story." If the story is wrong, then appropriate skepticism can be generated. Editors need to get such stories correct, as the consequences of legal action can be a major threat. Recall that it was the same newspaper that broke the Burt story that bought the Hitler diaries: "It was the day the thunder of *The Times* turned into a whimper and *The Sunday Times* was forced to sniff the stench of self-deceit" (see Kohn, 1986, p. 159). I note a more recent "scandal" breaking in a newspaper concerning another academic, Bruno Bettelheim (Bernstein, 1990). It is hoped that the accusers learn much from the case for and against Sir Cyril Burt before these accusations are repeated as if true. What should be discussed is the procedures for investigating these claims and the ease with which these claims can be made on the basis of such little data.

Regardless of the pros or cons of the case against Burt, there are many instances of ludicrous claims in secondary sources. Broad and Wade (1982) commented that Burt's work was "riddled with glaring statistical errors, later shown to be a sign of wide-scale fraud ... even more remarkably, the critics of Burt's position also missed the red flags that leaped out from his disputed writings" (p. 203). They contended that Burt was a cheat, who "used his mastery of statistics and gift of lucid exposition to bamboozle alike his bitterest detractors and those who acclaimed his greatness as a psychologist" (p. 204). They portrayed Burt as lonely and embattled, with a tormented imagination.

One often suggested way to avoid fraud is to make raw data available, although it is recognized how difficult this is. Most universities I have been in inevitably discussed the beauties of archiving, and may have written the appropriate computer code to efficiently and effectively archive staff and students' data. After some time it becomes evident that the few who use these data do not justify the large cost of entering it and maintaining it.

There are other ways of screening data. Psychology, like most academic disciplines, has a long-standing basis on the peer review system. Although it

is fallible (Fiske & Fogg, 1990; Mahoney, 1977; Peters & Ceci, 1980), it needs to be strengthened rather than replaced. Reviewers and editors need to be more accountable for errors (such as failure to recognize plagiarism), more attention needs to be placed on comments relating to dishonesty (provided the comments are reviewed with particular attention to evidence), and more attention and credit need to be given to replications (e.g., who will start *The Journal of Replications*). Perhaps psychology needs more critics such as exist in the arts, music, and literature and more journals such as *Contemporary Psychology*. The most recent effort to begin such a journal in education failed (*Contemporary Education Review*). If psychology and education had more formal mechanisms for excellence (e.g., Nobel awards) then more rigor might be placed on probity, although such rewards might encourage undesirable behaviors. The move to assessing academics on the basis of performance indicators can also have this effect, as many nonproductive academics may be forced to publish even if they lack original or worthwhile ideas of their own.

There is also a need for more biographies in psychology. Boring's biographies are must reading for all aspiring psychologists, and his beginnings need to be continued and placed in front of all. There needs to be caution, however, in psychobiography, which is more likely to use inductive post-hoc justifications for the behavior. The motives for fraud are many, and there is much difficulty in accepting either Hearnshaw's explanations for Burt's behavior or Joynson's explanations for Hearnshaw's. To claim that cheating "may be like cirrhosis of the integrity" or the result of a pathological personality akin to that of compulsive gamblers (Luria, 1975) ignores simpler explanations of greed, fame, or the need for consistency. Merely depending on the memories of many, still living colleagues about a former teacher, mentor, or leader are not the most reliable indicator unless these views are placed in perspective. We need to recall the Greek historian Thucydides' maxim: Men are more anxious to be called clever than honest and suspect honesty of simplicity.

Afterword

In the long run, Cyril Burt's evidence on heredity does not matter. Others have replicated his work, typically with the same results (Bouchard, 1987, Bouchard & McGue, 1981; Osborne, 1990; Rimland & Munsinger, 1977; see also Snyderman & Rothman, 1988). His contributions to education have been made and generally are now surpassed, his statistical work ignored or replaced by alternative procedures, and as Cronbach (1979) noted "he was hopelessly out of touch with mainstream psychology" (p. 1392). So today it is possible to forget him while not changing any cherished views. He appears, unfairly, to be relegated to the waxworks alongside other infamous characters who have been vilified, tried, but have no opportunity to reply. To question all his work is ludicrous (is the mathematics in *Factors of the Mind* a figment of his fraudulent efforts?). It is of little interest to determine exactly when Burt's work becomes suspect (Hearnshaw cites 1943; Clarke & Clarke, 1980a, 1980b and Kamin, 1974 say all of it), although there is probably merit in the case that Burt would have made a good career decision

by dying at age 60. The 13th stroke of the clock brings into question the previous 12, and most researchers are not using any of Burt's work to support their claims.

Yet at the time Burt was influential, if not preeminently influential in Britain and the colonies. For right or wrong, Burt deserves to be considered among the most influential psychologists of this century. His deviations, if substantiated, are serious, but minor considering his direct influence on education and psychology. He made major contributions to the backward, gifted, mental testing, use of statistics in educational psychology, child guidance, delinquency, military psychology, subsequent criticism of the 11+, and as Hearnshaw noted "Burt's main achievements were in the field of individual, or differential, psychology; in some branches of mathematical psychology; and in theoretical psychology of a humanistic type—all areas about whose status psychologists have as yet reached no consensus" (p. 319). He deserves to have his portrait hung alongside the other pioneers—Spearman, Pearson, Thomson, and Garnett. There is no doubt, however, that his alleged deviations will be his lasting marks on psychology. He is not Hearnshaw's "mad professor" or Joynson's "happy warrior," he is now the marginal psychologist.

The message for present-day psychologists is clear: exactness, scrupulousness, and attention to detail are needed. "Be honest; never manipulate data; be precise; be fair with regard to priority; be without bias with regard to data and ideas of your rival; do not make compromises in trying to solve a problem" (Mohr, 1979, p. 45). Scandal is a honeypot for many. It is astounding the number of academics who have passed judgment on Burt, who sneer at the views he espoused, but who have never read Hearnshaw, or now, Joynson. These two books are required reading to be informed about the debate; a careful reading of some of Burt's work is necessary; and both need to be tempered with the realization that Burt has had his influence on education and now has an influence on biographical and ethical standards in psychology.

Note

1. I thank J. Rogers and H. Swaminathan for their comments on this review.

References

Ammons, C. (1990). Joynson, R.B. The Burt affair. *Psychological Reports, 55*, 351.

Audley, R.J., & Rawles, R.E. (1990). On a defence of Professor Sir Cyril Burt. *The Psychologist, 3*, 360-361.

Beloff, H. (1990). Review of Joynson: The Burt affair. *British Journal of Psychology, 81*, 395-397.

Bernstein, R. (1990, November 4). Accusations of abuse haunt the legacy of Dr. Bruno Bettelheim. *New York Times*, p. E6.

Blinkhorn, S. (1989, August 10). Was Burt stitched up? *Nature, 340*, 439-440.

Bouchard, T.J. (1987). The hereditarian research program: Triumphs and tribulations. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy*. New York: Falmer.

Bouchard, T.J., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*, 1055-1059.

Broad, W.J., & Wade, N. (1982). *Betrayers of the truth*. New York: Simon & Schuster.

Burt, C.L. (1909). Experimental tests of general intelligence. *British Journal of Psychology, 3*, 94-177.

Burt, C.L. (1949). Alternative methods of factor analysis and their relations to Pearson's method of "principal axes." *British Journal of Psychology (statistical section), 2*, 98-121.
Burt, C.L. (1954). The differentiation of intellectual ability. *British Journal of Educational Psychology, 24*, 76-90.
Burt, C.L. (1955). The evidence for the concept of intelligence. *British Journal of Educational Psychology, 25*, 158-177.
Burt, C.L. (1966a). The genetic determination of differences in intelligence: A study of monozygotic twins reared together and apart. *British Journal of Psychology, 57*, 137-153.
Burt, C.L. (1966b). Parapsychology and its implications. *International Journal of Neuropsychiatry, 2*, 363-377.
Burt, C.L. (1968). An illustration of factor analysis. In H. Butcher (Ed.), *Human intelligence: Its nature and assessment* (pp. 66-71). London: Methuen.
Burt, C.L. (1969). Intelligence and heredity: Some common misconceptions. *Irish Journal of Education, 3*, 75-94.
Clarke, A., & Clarke, A. (1980a). Comments on Professor Hearnshaw's balance sheet on Burt. In H. Beloff (Ed.), *Supplement to the Bulletin of the British Psychological Society, 33*, 17-19.
Clarke, A., & Clarke, A. (1980b). Review of Hearnshaw (1979). *British Journal of Psychology, 71*, 172-173.
Clarke, A., & Clarke, A. (1990a). Review of Joynson: The Burt affair. *British Journal of Educational Psychology, 60*, 122-124.
Clarke, A., & Clarke, A. (1990b). The Burt affair. *The Psychologist, 3*, 74.
Conway, J. (1958). The inheritance of intelligence and its social implications. *British Journal of Statistical Psychology, 11*, 171-190.
Cronbach, L.J. (1979). Hearnshaw on Burt. *Science, 206*, 1392-1394.
Eysenck, H.J. (1990, September 2). Sensitive intelligence issues. *Spectator, 263*, 26-27.
Fancher, R. (1989, November 9). Fixing it for heredity. *London Review of Books, 11*, 19-20.
Fiske, D.W., & Fogg, L. (1990). But the reviewers are making different criticisms of my paper! Diversity and uniqueness in reviewer comments. *American Psychologist, 45*, 591-598.
Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
Gould, S.J. (1981). *The mismeasure of man*. New York: Norton.
Grafton, A. (1990). *Forgers and critics: Creativity and duplicity in Western scholarship*. Princeton, NJ: Princeton University Press.
Guilford, J.P. (1936). *Psychometric methods*. New York: McGraw-Hill.
Hearnshaw, L.S. (1979). *Cyril Burt: Psychologist*. Ithaca, NY: Cornell University Press.
Hearnshaw, L.S. (1990). The Burt affair: A rejoinder. *The Psychologist, 2*, 61-64.
Hudson, L. (1989, November 3-9). Recalling a scapegoat. *Times Literary Supplement*, pp. 1201-1202.
Jensen, A.R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39*, 1-123.
Jensen, A.R. (1974). Kinship correlations reported by Sir Cyril Burt. *Behavior Genetics, 4*, 1-28.
Joynson, R.B. (1989). *The Burt affair*. London: Routledge.
Joynson, R.B. (1990). The Burt affair: A reply. *The Psychologist, 2*, 65-68.
Kamin, L. (1974). *The science and politics of IQ*. Potomoc, NY: Erlbaum.
Kohn, A. (1986). *False prophets*. Oxford, UK: Blackwell.
Luria, S.A. (1975). What makes a scientist cheat? *Prism*, 16-18.
Macdonell, W.R. (1901). On criminal anthropometry and the identification of criminals. *Biometrika, 1*, 177-227.
Mahoney, M.J. (1977). Publication prejudice: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research, 1*, 161-175.
McAskie, M. (1978). Carelessness or fraud in Sir Cyril Burt's kinship data? A critique of Jensen's analysis. *American Psychologist, 33*, 496-508.
Mohr, M. (1979). The ethics of science. *International Science Reviews, 4*, 45.
Osborne, R.T. (1990). Cyril Burt's invariant kinship correlations: Another look. *The Mankind Quarterly, 31*, 163-170.
Pearson, K. (1901). On the systematic fitting of curves to observations and measurements. *Biometrika, 1*, 265-303.
Peters, D.P., & Ceci, S.J. (1980). A manuscript masquerade. *The Sciences, 35*, 16-19.

Plomin, R. (1989). Book review: The Burt affair. *Behavior Genetics, 19*, 739-740.
Rimland, B., & Munsinger, H. (1977). Burt's IQ data. *Science, 195*, 248.
Saint John-Roberts, I. (1976, November 26). Cheating in science. *New Scientist, 72*, pp. 465-469.
Schonemann, P.H. (1987). Jensen's g: Outmoded theories and unconquered frontiers. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy* (pp. 313-328). New York: Falmer.
Skanes, G.R. (1978). Correspondence. *Bulletin of the British Psychological Society, 31*, 201.
Snyderman, K., & Rothman, S. (1988). *The IQ controversy, the media and public policy*. New Brunswick, NJ: Transaction.
Steiger, J.H., & Schonemann, P. (1976). A history of factor indeterminacy. In S. Shye (Ed.), *Theory construction and data analysis in the social sciences* (pp. 136-178). San Francisco, CA: Jossey Bass.
Stein, D. (1989, August 24). The great IQ scandal. *Listener, 122*, p. 26.
Strachey, L. (1931). *Portraits in miniature*. London: Chatto & Windus.
Thieman, T.J. (1990). Burt. *Choice, 27*, 879.
Thurstone, L.L. (1947). *Multiple factor analysis*. Chicago, IL: University of Chicago Press.
Tucker, N. (1990, August 4). Mendacity or muddle? *Times Educational Supplement*, 19.
Tysoe, M. (1990, August 11). Burt the bogeyman. *New Statesman and Society, 2*, 30.
Valentine, C. (1965). Cyril Burt: A biographical sketch and appreciation. In C. Banks & P.L. Broadhurst (Eds.), *Stephanos: Studies in psychology present to Cyril Burt* (pp. 11-20). London: University of London.
Velicher, W.F., & Jackson, D.N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research, 25*, 1-28.
Zenderland, L. (1990). Burt again. *Science, 248*, 884-886.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 3, September 1991, 277-294*

PERSPECTIVES

Kohlberg's Legacy—More than Ideas
An Essay Review

WILLIAM J. HAGUE
University of Alberta

Kohlberg's contribution to psychology was not only scholarly but personal, according to the contributors to this most intriguing tribute to an eminent psychologist and educator. As an introduction to Kohlberg's ideas on moral development this book is clear, practical and helpful. However in addressing the topic of the helping professions, some intepretations are unnecessarily restricted, while others, particularly in the case of implications for counseling, seem to let admiration lead to exaggeration. This essay concludes with an examination of the convergence of Kohlberg's legacy with character and virtue psychology.

Kuhmerker, Lisa, (1991), *The Kohlberg Legacy for the Helping Professions*. Birmingham, AL: Religious Education Press, 234 pages, Paper Cover, $14.95 U.S.

Some people give their name to ideas. We switch readily from the adjectival "Freudian" to saying, "That's Freud"; we say, too, "This is Rogers." Lawrence Kohlberg has become a name that is an idea. When we talk about a particular stage of moral reasoning, it is convenient to say, "That's Kohlberg." This name has taken its place with those of Freud, Skinner, and Rogers as a turning point in the history of psychology. Agree with them or not, one cannot deny that each of these men sent psychology and education off in a major new direction. Kohlberg had much to do not only with reshaping psychological theory, but also with giving practical direction to education, especially moral education. He addressed a need of his time in the United States with a philosophy and an empirically based theory that purported to walk the narrow path of educating morally without indoctrination while maintaining a principled morality. Kohlberg seemed to be the answer to the problem of maintaining moral standards without recourse to religion or talk of moral virtues. Content thus yielded to process; and this fit well with the American political and educational landscape of the time. Not too many years ago Kohlberg's famous moral dilemmas and his quickly accepted stages of moral development seemed to offer salvation not only to moral and religious educators, but to teachers in general. Now, a book entitled *The Kohlberg Legacy for the Helping Professions* offers this same salvation not

Bill Hague is a professor in the Department of Educational Psychology where he coordinates and teaches the graduate counseling program at the master level. He also teaches graduate and undergraduate courses in the psychology of religious and moral development and education. He is author of the book *New Perspectives on Religious and Moral Development.*

only to teachers but to a wider audience of those who help. The authors maintain that Kohlberg's influence goes beyond education, and his "legacy" extends much more broadly than some had imagined. What legacy did Kohlberg leave to education and to the helping professions, and what legacy does this book convey? These are the major themes of this review.

A Labor of Love

First a word about the tone of Lisa Kuhmerker's book. It is a scholarly work in the tradition of academic objectivity. But the unsuspecting scholar must beware; this is something quite rare in academia; this book is also a labor of love, rich with subjectivity. It is lovingly written by Lisa Kuhmerker and others who were close associates and obvious admirers of Lawrence Kohlberg over many years. We can expect it to be delightfully biased. It is replete with vignettes and personal impressions of Kohlberg the man. The central focus in it is clearly Kohlberg himself and what *he* has to offer the helping professions. The name of the man, as we have seen, can stand for the theory; in this book it stands for the theory, but the personal aura of the man permeates the whole. It is not without significance that the cover of the book shows a photo of Kohlberg with the sun shining through his hair in a distinctly halo-like effect.

Works by and about Kohlberg are numerous. Many are philosophical, theoretical. Many are psychological, both theoretical and practical, aimed sometimes at furthering research in moral education; at other times, directed to schooling. The niche that Kuhmerker, who is editor of the *Moral Education Forum*, carves out for this book is both theoretical and practical for the educator, the counselor, and the clinician. This book, after presenting the theory and the research on which practice is to be built, attempts to expand our conception of the practical dimensions of Kohlberg, Kuhmerker notes (p. 2) that this expansion is possible because, while Kohlberg's original research was limited to young people, his longitudinal studies and the research of others revealed that adult reasoning too shows stages and can be fostered by the challenges of adult life. This, she says, has had "a profound influence on counseling, pastoral care, clinical practice and programs in corrective institutions."

Almost any introductory psychology text in the last 20 years that has touched on moral development has featured Kohlberg and has contained at least a table giving his stages of moral development (sometimes canonized as "*The* Stages of Moral Development"). Psychologists and educators have a love affair currently with stages of development. Researchers and teachers are happy to use stages as neat containers for categorizing people. There is, it seems, a tacit promise that if we know the stage we will know the person, how to teach the person, and what to expect of him or her. The Kohlberg bandwagon was jumped on by many who saw his contribution as just that—a stage theory of individual development one could grab quickly and use immediately. But there was on Kohlberg's bandwagon much more than "neat" stages for categorizing people. Beyond were Kohlberg's conviction and lifelong dedication to the belief that schools need a democratic structure not only to develop individual moral reasoning, but also to foster a just society at

large. The issue was bigger, as big as the man himself, and that is what Kuhmerker tries to capture in this book.

Kuhmerker's practical aim is clear from the beginning:

> Our primary goal is to present Kohlberg's work clearly, so that the practitioner can be helped in his or her professional activity. With a few exceptions we do not focus on Kohlberg's critics. The annotated list of suggested readings for chapters 2-4 guides the reader to some of these critical evaluations. The bibliography at the end of the volume encompasses every one of the references by Kohlberg and his associates. Selected references by other authors are included. A doctoral student or researcher will find in this volume all the information he or she needs to decide whether or not Kohlberg's work is relevant to a projected research topic. The educator or counselor can see clearly whether or not the work of Kohlberg and his associates has relevance for his or her particular role in the helping professions. (p. 4)

Thus she makes a fairly modest but helpful claim: not that this book will solve all your problems; only that it will give you enough information to make some decisions. Those who have plowed through Kohlberg's often tangled style, written it seems by someone in a hurry to get on to the next project, will appreciate this intention to condense and clarify. With the above statement Lisa Kuhmerker succinctly outlines the goals of this volume, modestly eschewing any claims to be publishing the encyclopedia of Kohlbergiana, but aiming instead at clarity and usefulness to students and practitioners. The way these practical goals are pursued and the success with which they are accomplished are the substance of this review.

Contents and Organization

The book is divided into two major sections, the first theoretical empirical and methodological, the second oriented toward the educational practitioner. Lisa Kuhmerker begins this work by coming from her strengths: "More than twenty years of involvement in the field of moral development and long-standing personal contact with Kohlberg himself and with colleagues in the Association for Moral Education" (p. 3).

Part I — Moral Development: Kohlberg's Theory, Research Findings, and Methodology

This volume on moral development theory, research, and education begins imaginatively and appropriately with Lawrence Kohlberg's own life story—*My Personal Search for Universal Morality.* It starts with his high school days, his problems with conventional morality and formal religion, and his attempts at resolution by reading Dostoevsky's *Brothers Karamazov.* It continues with his role as an ex-merchant mariner running the British blockade of Palestine. As a student he worked as an attendant in a mental hospital, "seeing from the bottom up both the needs of individuals and the problems of social justice in that arena" (p. 13). He studied under Bettleheim, Rogers, and Gewirtz, noting that, "None seemed adequate to deal with the patient's moral problems and the justice pattern of the institutions" (p. 13). An incident in the hospital in which he felt that a patient's just rights were violated provided the occasion to "refocus (his) energies."

Such personal accounts, far from being irrelevant to understanding theory and research, slice, I think, to the heart of the matter. If theories were the product of free-floating brains, we would not need to understand the people who conceive and research them. But, products of the whole person that they are, our understanding of theories is enriched deeply by knowing something of the human being from whom they came. Kuhmerker is astute in acknowledging this and acting on it for our benefit. The presence of the personal and subjective in her book is a richness few others give us.

From this personal framework for experiencing Kohlberg's legacy, Kuhmerker moves to "a systematic examination of his theory, methodology and research findings within the context of twentieth-century philosophy and psychology" (p. 4). Chapter 2 (written by Uwe Gielen) introduces Kohlberg's prototypical and now famous "Heinz dilemma"—a man faced with the choice of stealing a drug to save his wife's life. Gielen outlines seven basic key terms and basic assumptions of cognitive developmental theory applied to the moral domain. The chapter surveys the refining process Kohlberg and his associates went through and multiple revisions of assessment strategies that have resulted. This alerts readers to be critical of which version of the Moral Judgment Interview is being used in a particular piece of research. Gielen has done a service for the reader unfamiliar with Kohlberg's cognitive developmentalism. He begins (p. 18) with fundamentals, contrasting Kohlberg's ideas with the psychoanalytic approach of Freud and social learning theory as represented especially in the field of morals by Hartshorne and May (1928/1930). The role of this chapter seems to be that of a primer of cognitive developmentalism à la Kohlberg, including not only the standard table of Kohlberg's stages of moral development but a "pro and contra" table giving helpful examples of responses to the Heinz dilemma at nine stages and substages of moral development. Altogether Gielen's is a clear presentation of the basics, including a seven-term lexicon, a discussion of autonomy and heteronomy, and a summary of Kohlberg's Platonic philosophical foundations; all this in a lucid style invaluable to a Kohlberg novice.

Chapter 3 describes Kohlberg's major research—the 20-year longitudinal study of the moral reasoning of American males. This is followed by an overview of research studies under the headings: "cross-cultural," "gender differences," and "the relationship of moral judgment to action."

Chapter 4 endeavors to unravel a complex Gordian knot in Kohlberg's work—the scoring system he used to measure moral judgment. It is followed by a survey of two methods that flowed from Kohlberg's system: Gibbs's Social Reflection Method and Rest's Defining Issues Test, comparing the appropriateness of the three systems. Because learning to score Moral Judgment Interviews requires long, complex, and sometimes expensive training, this chapter, following the practical aims of this book, describes the rationale and the process so that prospective researchers may judge the appropriateness of the approach and the feasibility of using it.

This chapter is an opus in itself and probably contains more information than the typical student really needs. The casual reader, then, should not be discouraged if he or she feels information overload.

But, on the other hand, one should beware of oversimplifications. Notable among the statements in this detailed historical section is the following claim on *hermeneutic objectivism*:

> The coding guide successfully establishes a balance between the philosophical-psychological need for deep interpretation and the scientific need for objectivity and reliability. It is this balance that makes the coding guide such a striking example of objective hermeneutics. The tender minded approach of the humanities and the tough minded approach of the natural sciences are reconciled in Kohlberg's psychological work. (p. 65)

One would prefer a more modest statement.

At this point, the end of part one, Kuhmerker inserts a thoughtful little gift for the faithful reader: "Suggested readings for chapters 2-4." It is the kind of thing a helpful thesis supervisor would give a graduate student—an annotated bibliography filled not only with relevant references, but also with the kind of insider comments that come from an informed mentor who knows the ropes and can lead the student directly to relevant sources.

Part II—Kohlberg for the Educational Practitioner

Chapter 5, Kuhmerker says,

> serves as an introduction to chapters 6 to 8 on educational interventions with which Kohlberg was involved. It puts the significance and limitations of the use of hypothetical dilemma discussions into perspective. It alerts the reader to the fact that Kohlberg had an influence on the fostering of moral development through traditional curriculum ... and that much of the last ten years of Kohlberg's professional life were spent in active participation in the creation of "just community" alternative school programs. Chapter 5 explores why a school might wish to transform itself into a "just community" and what the complexities of implementing such a goal might be. (p. 6)

Chapter 6 explores the topic of dilemma discussions, giving "how to" information to help the teacher incorporate discussions into the curriculum. Chapter 7 returns to the biographical theme of Kohlberg's work in the schools of Massachusetts and New York, while chapter 8 explores the theme of the just community in the voice of one of its students in dialogue with Kohlberg. Chapter 9 pursues the topic of the just community still further by assessing the moral culture of the school as influenced by the "just community" experience.

Part III—Kohlberg for the Religious Education Practitioner

Chapter 10 looks at the major question of the influence of religious education on moral education. After indicating that Kohlberg's focus on structure rather than content may have opened the way for religious educators of various persuasions to find relevance in Kohlberg's work, the chapter focuses on the work of Fowler (1984) and Oser (1980) who have used Kohlberg's stages as a model for their own design of faith and religious development.

Part IV—Kohlberg for the Counseling and Clinical Practitioner

Chapter 11 moves on to issues of clinical diagnosis and intervention, beginning with a description of "competing counseling traditions" current at the

time Kohlberg began his work and making, as we shall see, some rather remarkable claims for Kohlberg's influence on counseling and psychotherapy.

Chapter 12 begins to bring the book to its conclusion, pointing out genealogical strains passing from Kohlberg through Selman, Kegan, and Noam.

Part V—The Kohlberg Legacy for the Helping Professions

Chapter 13 focuses the volume toward its conclusion, with James Rest providing a backward and forward look at Kohlberg's stage concept and cognitive construction of social morality. The work concludes in chapter 14 as it began—with something of a personal tribute to Kohlberg in the form of warm reminiscences shared at the award dinner of the Association for Moral Education when it honored Kohlberg in 1982. This, Kuhmerker (p. 7) says, is "the Kohlberg legacy to his friends ... If it had been possible, this is the way he would have wished to reach you—the educator, the counselor, the clinician, the practitioner who cares about the moral life."

Critique

The foregoing is largely an overview of what the book is about, how it is organized, its stated goals and general tone. Obviously it covers a range of topics too extensive to review each in depth. A reviewer must choose just some to look at and critique in detail. These should be chosen for their intrinsic importance and not merely for their interest to the reviewer or because he sees them as being somehow within his ken. Perhaps the following topics for detailed exploration are navigation points of this reviewer as he plots his course between Scylla and Charybdis.

It struck me while reading this book that the later Kohlberg is much more interesting than the earlier one. Perhaps this is partly due to the fact that what Kohlberg proposed in his initial years as revolutionary now is, in some quarters, old hat and part of the psychological establishment. But Kohlberg was a scholar always open to new ideas and ready at times to admit his mistakes, even if it took him a while to come around to it. Witness the article by Kohlberg and Candee, "The Relationship of Moral Judgment to Moral Action" in Kurtines and Gewirtz (1984, pp. 52-73) where Kohlberg picks up Frankena's distinction between "deontic" (principled) judgments and "aretaic" (responsible) judgments, and, although in 1984 he can't agree with Gilligan's proposition that there are two independent moralities, one of rules and justice and one of responsibility and care, he does go on to admit that high level moral judgments can be made using "intuition," "heart," and "conscience." These are remarkable words coming from one of the most cognitive of cognitivists. His willingness to move even further on this issue toward the end of his life is illustrated in his contribution to Modgil and Modgil's (1986) monumental work, *Lawrence Kohlberg, Consensus and Controversy*. Fascinating too, as we shall see later on, is Kohlberg's creation of an enigmatic Stage 7 when for much of his academic life he had claimed Stage 6 as the ultimate moral stage. This book, as we will see, misses some of the interesting and challenging later developments of Kohlberg by concentrating on an introductory approach and avoiding much of the criticism that lit

Kohlberg's intellectual fire. It is a pity that a book that unabashedly sings the praises of the man behind the theory neglects to show adequately his openness to change.

The Man Behind the Theory

Perhaps, then, the first major criticism of this book should concentrate on that which is at once both its strength and its weakness, on that which opens a door, but leaves it only partly ajar. By this I mean its innovative presentation of the man within the theory. As such it does the reader and psychology in general a real service by helping fill in the vacuum that so many academic publications leave: What kind of a person is this who proposes these great theories? Are they the result of sheer cerebral activity? Or do the ideas not flow from who the person is, what he has experienced, and, perhaps most of all, what he has suffered in his life? The anecdotes in this volume are mostly in the absent-minded, lovable, dedicated professor area. As such they add color and warmth and lead the reader to greater depths of appreciation of the man, his humor, his dedication, and his ideas. But more could be done with this. Years ago, it seems to me, something very different began to come from the pen of a man who had written his cognitivism in abstract and sometimes muddied prose. It came out in poetic, mystical words that spoke about a dimension beyond mere reason. Kohlberg (1970) in simple, photocopied class notes, entitled *The Ethical Life, the Contemplative Life and Ultimate Religion—Notes Toward Stage 7* presented a deep glimpse of himself that was more obscured in his formal publications. He begins (p. 2) by replying to his students' two objections to his moral stages: their rational intellectualistic quality and their closed quality. "While the stages are closed," he says,

> much of the appeal of the concept of development is the appeal to an open and individual process based on personal experience not conceptual thought. I have tried to convince you that a solution to moral problems must ultimately be a closed and permanent solution; it must be universal and it must be rational. If it is not, you will be back in the quagmire of relativism.

"But," he continues,

> I appreciate your demand for a concept of development which is more open, more vague and individualistic, and more dependent upon personal, non-rational experience. So in talking to you of my Stage 7, I will not be ... drawing upon the authority of scientific findings and moral logic ... I will only be talking to you about my tentative efforts and those of others to make sense of experience, once you are aware of Stage 6 morality and scientific psychology ... I hope to take away from stage 6 excessive demands placed on it, the demand that it provide the answer to the problems of the meaning of life. (pp. 2-3)

He goes on (p. 3) to speak about "the logic of contemplation and "mystical logic." He says, (pp. 3-4) "Ultimately we all know that the deepest feeling is love, and the ultimate reality is life, and our problem for today is how we can love a life which makes burdensome demands upon us, and yet does not even console us by telling us the world is moral or just." To address these questions of ultimate meaning the "logic of action" encompassed by his six stages of moral reasoning fails because "life itself is not a rational process" (p. 21). Our

birth, our life, and our death are processes beyond reason, so he postulates a metaphorical Stage 7, which is religious in its mystical experiential, meaning-making dimensions.

> The pains of life which religion hopes to console are the pains of death, of unrequited love and separation and of injustice, and it attempts to console them by postulating the reality of something ideal or infantile. These pains, religion claims, reflect the discrepancy between our actual existence and some concepts of ideal or infinite limits in terms of which our existence falls short.
> To be pained by death is to conceive and value immortality,
> to be pained by injustice is to conceive and value ideal justice,
> to be pained by lack of love is to conceive of ideal love.
> To be pained by the finite and non-ideal is to recognize the reality of the ideal and infinite as something more than a relative human construction.
> If the world's injustice were nothing more than one's hedonistic pains, it would not be crushing. If the concept of eternity were nothing more than a fiction, death would be no pain. It is, then, human ideals which cause the existential pains, and their power is, in a sense, a sign of their reality.
> (1970, p. 36)

Kohlberg wrote this passage as prose, but I have taken the liberty of breaking the central lines of parallel structure into poetic lines because this passage, unlike so much of Kohlberg's scientific writing, sings like poetry. These words are too potent and too beautifully put to be just ideas. These are, in heightened contrast to scientific terminology, words of poetry and personal perception; one can intuit deep pain, tracing the inner journey of a person who has himself searched for ultimate meanings. One can pick up from Kohlberg's Stage 7 themes that these transcendent dimensions are not just the result of thinking, but of experiencing, and painful experiencing at that.

Kuhmerker includes a one-page allusion to Stage 7 in her chapter on religious education. But it takes a rather matter-of-fact, cooler approach to the topic. The references she quotes (Kohlberg, 1984) are Kohlberg's more formal, scientific, and less personal presentations on the topic. But there is too much in Kohlberg's willingness (or rather need) to break with sheer rationality and go beyond to deeper personal, religious dimensions to treat this as a mere interesting offshoot of a grand, rational theory. I think it tells us much about the bigness of the man open to see other possibilities, the depths of the man to experience new understandings, and the sensitivity of the man to know the reality of pain. One wonders why Kuhmerker does not recount this important dimension. Surely it is an important part of the Kohlberg legacy.

My own sparse encounters with Kohlberg have perhaps something to offer here. I attended the moral education conference in Leicester, England at which Kohlberg made one of his first personal forays into the British Isles. One afternoon, following his perennial critic, R.S. Peters, he spoke to the philosophers and educators in attendance. Judging by the mumblings in the audience, he did not in this initial presentation make a very positive impression. But that evening he gave a special extra lecture on his Stage 7. This lecture was much more personal and flowed almost lyrically. I was reminded

of the photocopied lecture notes (quoted above). After the lecture, I stopped him in the dining hall, prefacing my question with the eternal student lead: "May I ask you a personal question...?" To this I got the reply, "Try me." My question said in effect, "It seems to me that your Stage 7 is not just an academic adjunct, an afterthought tacked onto the classic cognitive stages, but it comes from your own personal religious experience." His reply was prompt and straightforward, briefly alluding to his return to the religion of his youth. His brief account of his own spiritual quest, given simply and openly there in the refectory of Leicester University, confirmed my intuition that such a profound appreciation of the transcendent dimensions of the person had been born not in the theorizing of academia, but in the disintegration and rebirth of real life. We talked for a while about the beauty that usually surrounds these mystical experiences, until someone interrupted with more mundane matters. One has to admire such personal integrity, depth, and sensitivity.

Gender Differences

Context is important here. Kohlberg was much criticized in his lifetime, notably by one of his former students Carol Gilligan for having concentrated his research and his conclusions on male subjects, favoring a cognitive approach deliberately abstracted from individual relationships in the interests, it seemed, of moral objectivity. Gilligan (1982) asserted a need to listen to the voices of women who, she claimed, tended to solve moral problems more in terms of caring and responsibility. The result was an appreciation of broader, personal dimensions of moral thinking. But there was danger that two camps would be built in the moral desert—one an ethic of principle, the other an ethic of care; one stereotyped as "male," the other as "female." Despite Gilligan's assertion that the two were complementary, the way was opened for opposition and exclusivism. Thus gender differences became a popular and hot topic.

Gielen, in the course of his review of research on moral reasoning (chapter 3) approaches this prickly question of gender differences. Consistent with his lucid style, he contrasts Kohlberg and Gilligan in these terms:

> While Gilligan's theory focuses on a broadly defined "moral self" Kohlberg's deontic conception of moral reasoning is restricted to moral decision making in situations involving the clash of competing moral claims ... In Gilligan's theory there exists a distinct danger that processes of moral judgment are assimilated to, and confused with personality organization.

It is interesting to note that it is all right for Kohlberg to be restricted to a deontic conception, but there is a "danger" in Gilligan looking at moral judgment in holistic personality terms. Are the disciples more orthodox than their master?

But Gielen must be given credit for the openness of his appreciation of his mentor. In the conclusion of chapter 3 he says:

> Kohlberg's earlier tendency to equate the moral domain with justice reasoning may have induced tunnel vision in some of his followers, but it kept his early theoretical and methodological efforts sharply focused on a fundamental dimension of morality. In his later years, he expanded his vision to include

> concerns for benevolence, beneficence, caring and responsibility, although his drive toward a broader conception of the ethical self was cut short by his death. The expansion of his moral vision is most salient in his educational efforts ... Without justice there could be no true morality for him. Caring and benevolence were important considerations in his own personal life, and they remained important for his interpersonal theory of morality. But he insisted that while care can not be expanded to include everybody—a commitment to justice can. (p. 59)

There is a risk that the ethic of care will be identified merely as a "feminist" approach to making moral judgments—a conclusion that would be constricting to the very notion of care as something *all* human beings should be capable of. Perhaps the task is not, in the interests of orthodoxy, to maintain Kohlberg's earlier, more restricted views, but to appreciate the more holistic approach he came to appreciate more, and to which the ethic of care and responsibility invites us. Perhaps the call is not to keeping orthodox guard, but to advancing into new ways of achieving moral objectivity that include but go beyond cognitivism.

Gielen's contributions to this book are generally strong, clear, and positive. They sustain Kuhmerker's aim to be maximally helpful.

One could go on at this point with a critique of those parts of Kuhmerker's book that extol the merits of Kohlberg's legacy for education in general and moral education in particular. But sufficient ink and paper have already been spent on this over the years, and further surveys of the pros and cons may be redundant. Instead I would like to look at those sections of the book under review that forge more into the hinterland of Kohlberg's legacy to the helping professions: notably religious education and counseling. It is appropriate here, while returning to the context of the whole book and before leaving the theme of gender issues, to note that in this volume nursing is conspicuous by its absence from major consideration as a helping profession. Do the authors feel that Kohlberg has no noteworthy legacy to offer this profession?

Kohlberg and Religious Education

Chapter 10 of this volume focuses in on the influence of religious education on moral development, attempting to dissect and analyze that sensitive area where religion and moral education intersect. This chapter does much to clarify the issues and underline the problems of Kohlberg and religion. But as I read Kuhmerker's contribution here, I see her using the word *religion* only in its formal, institutional sense, ignoring the distinction between that kind of religion and religion as personal and communal experience of the numinous, the transcendent—the kind of religion James (1958) explored in his *Varieties of Religious Experience,* and Jung (1938) celebrated throughout his lifetime. With the limited notion Kuhmerker utilizes, much of what she has to say does follow logically. But without the fuller understanding of the richness of religion, much of what could be said about the relation of religion and morality is lost.

Kuhmerker comes close to using this distinction under the aegis of Lee when she quotes him as saying:

> Kohlberg operationally defines religion as religious affiliation. Thus his data do not show that morality is autonomous from religion but from religious affiliation (Lee, 1980, p. 33). Furthermore Lee points out that if Kohlberg states that activities engaged in by the self are developmental, and "religion is actively engaged in by the self" then religion is developmental and can be investigated directly or indirectly in all its dimensions. (Lee, 1980, p. 336) (p. 165)

But instead of picking up this crucial distinction Lee has laid out in the clearest of terms, Kuhmerker chooses to ignore it on experimental grounds: "One might well ask however how Kohlberg would have studied the influence of religious education, other than to have correlated it with the process of being religiously affiliated" (p. 165). She then slips back into her overly simple dichotomies of religious versus secular by saying: "To look for other causes brings us back once more to assumptions of the 'extrinsic zap of grace' or the secular Piagetian-Kohlbergian assertion that the developing self is the constructor of its own philosophy" (p. 165).

This failure to see religion as developmental, something in process, and to identify it instead almost solely with content leads Kuhmerker to shift words when she looks at religion as process and call it "faith development" in the light of Fowler's (1984) work. Even the word *faith* begs for clarification it does not get in this book; *faith* can range in meaning from the broad process concepts of Fowler (1984) and Tillich (1958) as a search for ultimate purpose and meaning in life to a most conservative view that "the faith" is a collection of specific tenets to be learned and held to, unchanging for a lifetime.

Overall, Kuhmerker, following the tenor of this volume, presents the potential religious educator with a rather concise summary of the interaction of Kohlberg's concepts of moral education and a variety of important names and approaches to religious education. For beginning religious educators this chapter nicely surveys the relevant Christian and Jewish religious educators. As a short cut to important names like Fowler, Oser, Lee, Groome, and Chazan, this chapter is a précis of introduction to some contemporary religious education thinking. It concludes with a typically Kuhmerkian device—six points of practical advice entitled: "Guidelines for Parents and Educators Who Want to Foster Religious Awareness and Commitment" (pp. 168-169). They are based on the premise that fostering moral reasoning "enables children to reason about all issues, including religious issues, in increasingly mature ways" (p. 168). One cannot quarrel with the importance of encouraging reasoning in moral education, particularly when some attempts at "religious" education are replete with examples of deliberately ignoring reason. But the theme here is Kohlberg's "reasoning is enough." One wonders if this is a real reflection of the later Kohlberg, given the broader view of moral choosing he began to take particularly in regard to religious issues (Kohlberg, 1974, 1984; Kohlberg & Candee, 1984).

Kuhmerker emphasizes (p. 157) that, "It is the public school's role to foster the capacity to reason about moral issues without 'teaching' the belief system of any organized group." This is true. But I think there is a danger of presuming that the obverse side of this coin is also true: that the religious school is always and everywhere interested only in teaching its own belief

system. There have been centuries of traditional catechising that has meant "conversion," but recent statements by some mainline churches show a promising emergence from a narrow position of rightness to a more ecumenical and humanistic concept of the whole person in process. Statements of ecclesiastical purpose are not, of course, necessarily followed by universal implementation, but they do point to the error of presuming that religious education must of necessity be indoctrination in narrow content. Reasoning can be and is encouraged in some religious settings with or without Kohlberg. Indeed, the religious setting is the most suitable situation for reasoning to flourish in the context of the whole person. Kohlberg himself (1974) was well aware of that.

Hyde (1990) has written an encyclopedic survey of empirical research in religion in childhood and adolescence, including an entire chapter on religion and morality in adolescence. It is interesting that in the over 1,800 references in the book, not one acknowledges Kohlberg's research. Where his name does appear, it is a passing reference to his foundational philosophy linked with Piaget or his influence on Oser's stages of religious development. Perhaps this says something about Hyde's British origins or about Kohlberg's North American popularity. Such an omission is, I think, regrettable, failing to recognize the monumental contribution Kohlberg with all his research flaws and sometimes questionable philosophy actually did make to education, moral education, and indeed to religious education.

Kohlberg's "Contribution" to Counseling

Richard Hayes authored chapter 11, "Counseling and Clinical Implications of Lawrence Kohlberg's Developmental Psychology." At this point in the book, admiration for the seeming universal importance of Kohlberg seems to take over from objective fact. From the introduction, the unwary reader might conclude that 1969 marked the dawning of a new age of counseling when that profession, thanks to the publication of a Kohlberg article (presumably the "Stage and Sequence" paper in Goslin, 1969, although that is not clear) suddenly became capable of a "full appreciation" (p. 173) of a developmental view.

One claim leads to another and we have Hayes (p. 173) generalizing that, "Prior to Kohlberg's seminal work in moral development counselors generally accepted one or the other of two competing views of human development." (Already it sounds too simple.) Hayes then explains that Kohlberg distinguished the "cultural transmission model" from the "romantic model." As someone who was a counselor educator in the 1960s and 1970s, I tried to locate myself in this dichotomy. Caught on the horns of this dilemma, I felt less comfortable on the behavioristic horn and so found myself labeled a romantic. Hayes then switches to his own terminology (without warning the reader who must fish in the words for his or her own continuity) to a bifurcation labeled "environmentalist" and "maturationist." Even though I profess a concern for the environment, I thought that I probably belonged in the maturationist category. And that seemed to fit until I came to the words: "For the maturationist, change is attributed to growth and differentiation. Environmental factors are recognized, but only as *placing limits* [emphasis

mine] on individual development, not as directing its course" (p. 174). There I felt I must part company with this limiting "maturationist" view of development and whatever it had to say about counseling.

To add to the confusion, Hayes then goes back to Kohlberg's original "cultural transmission" versus "romantic" terminology, presumably to prepare the reader for quotes from the master. These quotes from Kohlberg and Mayer (1972) represent the thinking of Kohlberg almost two decades ago; they do not reflect the sophistication of thinking and distinctions that existed in counseling theory 20 years ago, never mind today! For example, Hayes (p. 175) quotes statements from Kohlberg describing the counselor as "value-free" or "value-neutral." It is not clear from the quotations at this point whether Kohlberg is endorsing the idea or not; but later we will find that Kohlberg appropriately denounced value neutrality in counseling. Hayes is giving us a watered-down and confusing Kohlberg.

It is remarkable how Hayes gets into the value neutral problem in the first place. He does it by claiming that both the romantic and industrial psychology strategy of the cultural transmissionists come from a reductionism that betrays what Kohlberg has called a "bag of virtues" (p. 175). I cannot see how people who profess the teaching of specific virtues can have much to do with value neutrality except by way of contrast.

Another counseling issue Hayes touches on is the role of crisis in development. He notes Kohlberg's recognition of critical periods in development. But he then goes on to say in the next sentence (slipping into a narrower meaning of the word *crisis*) that, in the maturationist view of counseling, "Treatment is reduced to crisis intervention or the prevention of future crises." Again this may have been prevalent thinking 20 years ago, to which period Hayes is regrettably continually taking us back in order to show the implications Kohlberg for counseling today. But today we see a crisis as an opportunity, something to be grasped and used for growth, a disintegration of a lower level so that it can be replaced by higher level functioning (Dabrowski & Piechowski, 1977). Rather than these stale statements from the distant past, the theme that moral development comes about not *despite* but *through* moral crises (Hague, 1976) is much preferable, more appropriate, and a more faithful expression of the radical Kohlberg legacy.

Finally (p. 176), Hayes tells us that neither of these two categories, the behaviorist or the romantic, to which counselors of a quarter century ago paid their allegiance is the right way. Kohlberg has instead offered the "tradition of progressivism which encourages the nourishment of the individual's natural interaction with a developing society or environment, and a cognitive-developmental psychology to explain the nature of human development" (p. 176). In the section on basic assumptions, a naive reader may wonder at the genius of Kohlberg, who is purported to have constructed an elaborate theory of development that seems to incorporate the best of existentialism's meaning-making, behaviorism's social constructivism, symbolic interactionism, and organismic theories' attention to qualitative reorganization in the tradition of Werner and Kaplan. At the summit of all this, Kohlberg emerges, of course, with a cognitivist Piagetian approach, and does

so in a progressive tradition, following Dewey. It is a bit much for anyone to synthesize.

Instead of the elaborate hermeneutics Hayes uses to dig all these facets of Kohlberg primarily from the depths of a single paper, one would appreciate more an elaboration of a few key points in Kohlberg's psychology and what counselors can learn from them. For example, role taking is surely a major theme of Kohlberg's psychological and educational thinking. It is essential to development in the moral atmosphere of the school. The counselor has in his or her vocabulary the word *empathy* and he or she may spend a lifetime trying to get it right so that there is involvement with a client without getting all mixed up in the client's problem. If a researcher could find that Kohlberg had something to say about this empathy in the context of his experience with role-taking, he or she could indeed make a fine, *original* contribution to counseling.

Would it be too simple to say that Kohlberg was a moral philosopher and educator, and that consequently he may have something to tell us about the role of moral conflict in counseling? Clients frequently have moral problems mixed in with other personal concerns. Would it be simplistic to suggest that Kohlberg's legacy is to insist that, instead of skirting moral issues, one must confront them? Then to say that this confrontation, though painful, can be growthful? And then to say that Kohlberg gives us at least one practical way of handling personal issues—with reason?

Even more basically, one could take Kohlberg's fundamental cognitivism and find it akin to Ellis's Rational Emotive Therapy, but with a dimension that Ellis does not seem to have: an appreciation that thinking is not just rational or irrational, but that reasoning itself is multileveled. A client would, with this appreciation, be encouraged not just to think rationally, but to think at a higher level. The concept of moral levels would be a major contribution to Rational Emotive Therapy that has been wanting in this area.

After reading this chapter, one is not as convinced, as the author seems to be, of the great innovative insights into counseling offered by Kohlberg, especially when so much of it depends on a single article by Kohlberg and Mayer published in 1972. It shows, instead, that Kohlberg understood to some extent what was going on in counseling at the time, pigeonholed counseling in rather generic categories, and presented his own alternative approach as part of the progressive tradition of Dewey in the cognitive developmental framework of Piaget. Altogether, this chapter is a tribute to Kohlberg's broadness of view, his scholarship, and his personal concern for people and for counseling (Kohlberg, 1975; Kohlberg & Wasserman, 1980). But he is far from being the originator of any special school of counseling. The most that can be said is that his assumptions can support and supplement those of some others who have better worked out and more elegant theories of counseling.

A Question of Virtue

Kohlberg's early writings (1958, 1963, 1964) were highly critical of Hartshorne and May's (1928/1930) landmark research on character education. They had tried to assess the success of moral education in terms of

virtues practiced in real-life settings. Little was empirically supported. The shakiness of this research led Kohlberg, in his concern for an objective morality, to shun traditional character education, loaded as it was with content, and to look rather in the field of cognitive process. The term *bag of virtues* took on a scornful note. To right the imbalance of former years of character education (often in a religious setting) Kohlberg focused deliberately on the structure of reasoning about social justice rather than its content. This was historically opportune and suited the needs of public education in the United States. It fit well with progressivism and a democratic society eager to encourage individual self determination, offering at the same time the objectivity of principled behavior founded now not on many virtues, but on one—justice.

This process emphasis promised to avoid the indoctrination of content identified largely with private and religious schools. Interestingly, Kuhmerker (p. 157) even claims that this emphasis on process has recently come full circle to make Kohlberg's work "relevant to theologians and religious educators of widely divergent points of view." But behind this thinking is an assumption about religious education—that it is invariably obsessed with content, and must come full circle, back on its own tail so to speak, to appreciate the role of process in religious education. Such stereotypes are far from the reality of contemporary religious education which is aware of the importance of a happy combination of process and content. Content is an important part of moral education, and moral education is in a healthier condition when content is part of the open curriculum rather than buried in a hidden curriculum. In political situations demanding a split between church and state, some moral education approaches have become preoccupied with freeing themselves from content in an effort not to be "religious." But to say that moral education can be truly content-free is tantamount to the protestations of psychologists a few years ago that their counseling must be value-free. We have grown enough in the field of counseling to admit that values are present and active. It is time in moral education to admit that virtues are taught, consciously or unconsciously, and their practice encouraged as good habits of behavior. Character education has not for the last 20 years been the "in" terminology of moral education. Concern for an objective morality has concentrated on a Kantian categorical imperative and the role of reason in discovering a universal justice. Kohlberg has been the eminent champion of this cause. But recently other approaches are being explored. One can see it in the writings of Kupperman (1988) and Nussbaum (1988) who wish to develop the Aristotelian tradition of the possibility of moral objectivity in a world that recognizes a multiplicity of virtues. *Character education* and *virtue* are words coming over the horizon, not as the dark monsters they have in the past been taken for, but as a sunrise, proclaiming a new day in moral education (Hague, 1990).

Perhaps it seems strange to juxtapose *new* and *virtue*, for is virtue not an ancient word encrusted with medieval ideas of the indoctrination of God's will? Is it not contrary to all the autonomous rationality Kohlberg challenged? MacIntyre's *After Virtue* (1981) seems to lament the death of virtue. But authors such as Meilander (1984) and Barden (1990) have reminded us

that virtue is, after all, the moral strength of a good habit. If we can get away from seeing habit as a low-level, knee-jerk type response and perceive it instead as something consciously learned, developed, and strengthened with real-life practice, it may well serve again as the idiom for moral education. The work of Hauerwas (1985) has promised a new bridge between psychology and theological ethics. Important social commentators such as Lasch (1991) are drawing our attention to the need of specific virtues in our time, particularly the need for strong hope to replace shallow optimism.

Gielen (p. 49), as we have seen, warned us of the "danger" of "confusing" moral judgment with personality organization. Dykstra and Parks (1986) do not seem to share this fear:

> Neither Piaget nor Kohlberg intends to provide a theory of ego or personality development. Both, therefore, have approached the task of identifying the forms of reason or logic characteristic of different "stages" in human thought without making a critically important distinction: They have not attended to the differences between constitutive knowing in which *the identity or worth of the person is not directly at stake* and constitutive knowing in which it is. This has meant that Kohlberg has avoided developing a theory of the moral self, of character, or of conscience. Strictly speaking, his stages describe a succession of integrated structures of moral logic. He has given very little attention to the fact that we "build" ourselves through choices and moral (self-defining) commitments. His theory, for understandable theoretical and historical-practical reasons, has not explicated the dynamics of the inner dialogue in moral choice between actual and possible selves. (p. 22)

The task of the post-Kohlbergian moral educator will be to explore the problems and the possibilities of character education, how virtue might be taught and practiced, and how we might evaluate it with objectivity. This, interestingly, can be done in keeping with those concerns of the Kohlberg legacy which truly are most basic.

Conclusion

What, then, can be said of Kohlberg's legacy and Kuhmerker's efforts to pass it on? That Kohlberg left a rich legacy to pass on to psychology and moral education needs no further proof than this volume provides in the act of passing that legacy on. If one wishes to explore the fundamentals of Kohlberg's moral development from theory to application in the classroom and research, this book is probably the clearest, most comprehensive, and helpful publication in existence. Likewise, if one is interested in a glimpse at the man behind the theories, this book will provide that insight with a little help from his friends.

If one is interested in the implications of Kohlberg for religious education, this book is helpful and interesting. But, given the restricted parameters imposed on this field by its authors, one needs to be wary. However, if one is interested in exploring the implications of Kohlberg for the helping professions of counseling and psychotherapy, the sections of this volume devoted to that topic are hopelessly inadequate and out of date and naively distort the Kohlberg legacy out of all proportion. Kohlberg deserves a more balanced appreciation of how his ideas can be incorporated into the more sophisticated

theories of counseling that do exist and could benefit from his basic insights without pretense that this part of his very real legacy is more than it really is.

If one is only interested in Kohlberg himself, what kind of a man he was, this book is probably the best there is. But this must be taken with the proviso of a certain bias because it is written lovingly by friends and associates who hold his name in benediction.

References

Barden, G. (1990). *After principles.* Notre Dame, IN: University of Notre Dame Press.

Dabrowski, K., & Piechowski, M. (1977). *Theory of levels of emotional development* (2 vols.). Oceanside, NY: Dabor.

Dykstra, C., & Parks, S. (Eds.). (1986). *Faith development and Fowler.* Birmingham AL: Religious Education Press.

Fowler, J. (1984). *Becoming adult, becoming Christian.* San Francisco, CA: Harper & Row.

Gilligan, C. (1982). *In a different voice.* Cambridge MA: Harvard University Press.

Goslin, D. (Ed.). (1969). *Handbook of socialization theory and research.* Chicago, IL: Rand McNally.

Hague, W. (1976). Positive disintegration and moral education. *Journal of Moral Education, 5,* 231-240.

Hague, W. (1990). Shared values for shared life meaning: The transcendent dimension of the global village. In B. Holyst (Ed.), *Mental health in a changing world* (pp. 27-38). *Warsaw: Polish Society for Mental Health.*

Hartshorne, H., & May, M.A. (1928/1930). *Studies in the nature of character, Vol. 1: Studies in deceit; Vol. 2: Studies in service and self-control; Vol 3: Studies in organization of character.* New York: Macmillan.

Hauerwas, S. (1985). *Character and the Christian life.* San Antonio, TX: Trinity University Press.

Hyde, K. (1990). *Religion in childhood and adolescence.* Birmingham AL: Religious Education Press.

James, W. (1958). *Varieties of religious experience.* New York: New American Library.

Jung, C.G. (1938). *Psychology and religion.* New Haven, CT: Yale University Press.

Kohlberg, L. (1958). *The development of modes of thinking and choices in the years from 10 to 16.* Unpublished doctoral dissertation, University of Chicago.

Kohlberg, L. (1963). Moral development and identification. In H. Stevenson (Ed.), *Child psychology. 62nd yearbook of the National Society for the Study of Education* (pp. 277-332). Chicago, IL: University of Chicago Press.

Kohlberg, L. (1964). The development of moral character and ideology. In M.L. Hoffman & L.W. Hoffman (Eds.), *Review of child development research* (vol. 1, pp. 383-431). New York: Russel Sage Foundation.

Kohlberg, L. (1970). *The ethical life, the contemplative life and ultimate religion—Notes toward stage 7.* Photocopied notes for Social Sciences 154.

Kohlberg, L. (1974). Education, moral development and faith. *Moral Education Forum, 4,* 1.

Kohlberg, L. (1975). Counseling and counselor education: A developmental approach. *Counselor Education and Supervision, 14,* 250-256.

Kohlberg, L. (1984) *Essays on moral development, Vol. II: the psychology of moral development* (pp. 52-73). New York: Harper & Row.

Kohlberg, L., & Candee, D. (1984). The relation of moral judgment to moral action. In W. Kurtines & J. Gewirtz (Eds.), *Morality, moral behavior and moral development.* New York: Wiley.

Kohlberg, L., & Mayer, R. (1972). Development as the aim of education. *Harvard Educational Review, 42,* 449-96.

Kohlberg, L., & Wasserman, E. (1980). The cognitive-developmental approach and the practicing counselor: An opportunity for counselors to rethink their roles. *Personnel and Guidance Journal, 58,* 559-567.

Kupperman, J. (1988). Character and ethical theory. In P. French, T. Uehling, & H. Wettstein (Eds.), *Midwest studies in philosophy, Vol. XIII: Ethical theory: Character and virtue* (pp. 115-125). Notre Dame IN: University of Notre Dame Press.

Kurtines, W., & Gewirtz, J. (Eds.). (1984). *Morality, moral behavior and moral development.* New York: Wiley.

Lasch, C. (1991). *The true and only heaven.* New York: Norton.

Lee, J.M. (1980). Christian religious education and moral development. In B. Munsey (Ed.), *Moral development, moral education and Kohlberg.* Birmingham AL: Religious Education Press.

MacIntyre, A. (1981). *After virtue.* Notre Dame IN: University of Notre Dame Press.

Meilander, G.C. (1984). *The theory and practice of virtue.* Notre Dame, IN: University of Notre Dame Press.

Modgil, S., & Modgil, C. (1986). *Lawrence Kohlberg: Consensus and controversy.* Philadelphia, PA: Falmer.

Nussbaum, M. (1988). Non-relative virtues: An Aristotelian approach. In P. French, T. Uehling, & H. Wettstein (Eds.), *Midwest studies in philosophy, Vol. XIII: Ethical theory: Character and virtue* (pp. 32-53). Notre Dame IN: University of Notre Dame Press.

Oser, F. (1980). Stages of religious judgment. In C. Brussel (Ed.), *Toward moral and religious maturity.* Morristown NJ: Silver Burdett.

Tillich, P. (1958). *Dynamics of faith.* New York: Harper & Row.

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format

- Manuscript in triplicate.
- Typewritten or printed, double-spaced throughout.
- Abstract of approximately 100 words on a separate sheet.
- Manuscript not to exceed 6,500 words excluding graphics.
- Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.

Referencing

- Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s).
- Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format.
- All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
- Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes*; any citations in notes follow the same format as other references.

Graphics

- Each table and figure on a separate page at the end of the manuscript.
- Figures, charts, and diagrams in camera-ready form.
- Number tables in Arabic numerals and center over each heading, e.g., *Table 1.*
- Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1: ...*
- Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*

Style

- *AJER*'s editorial style conforms closely to the *Publication Manual of the American Psychological Association.*
- Editorial changes may be made to manuscripts.
- For spelling, consult *Webster's New Collegiate Dictionary;* spelling in quoted material remains as in the original.

Diskettes

If manuscript is accepted for publication and has been prepared using a Macintosh, IBM, or IBM compatible microcomputer, please send diskette. Microsoft Word, MultiMate, WordPerfect, WordStar, Xerox Writer, XyWrite, and DisplayWrite 3 only are acceptable.
Important: Please indicate program used.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index*; appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*



THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXVII NUMBER 4 DECEMBER 1991

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Robert H. Short

Consulting Editors

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

Janet Donald
McGill University

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Lilian Katz
University of Illinois

Les McLean
Ontario Institute for Studies in Education

John McLeod
University of Saskatchewan

Linda Phillips
Memorial University of Newfoundland

Lauran Sandals
University of Calgary

Keith Sullivan
Dalhousie University

Phil Winne
Simon Fraser University

Ruth Wright
University of Ottawa

Faculty Advisory Committee

Harvey Zingle (Chair)
Fern Snart
Erwin Miklos
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Second Class Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXVII, Number 4, December 1991

Special Edition:
50th Anniversary of the
Faculty of Education

HARVEY ZINGLE	295	Editorial: You've Come a Long Way!
ERWIN MIKLOS	297	Trends in Doctoral Research in Educational Administration
HENRY W. HODYSH	323	A Note on History, Educational Policy, and the Uses of the Past
GRACE V. MALICKY	333	Myths and Assumptions of Literacy Education
DAIYO SAWADA	349	Deconstructing Reflection
HEIDI KASS	367	The Transformation of Curricular Knowledge
ROBERT F. MULCAHY	385	Developing Autonomous Learners

ISSN 0002-4805

The Alberta Journal of Educational Research *Vol. XXXVII, No. 4, December 1991, 295-296*

EDITORIAL

You've Come a Long Way!

I am delighted as the eighth dean of the Faculty of Education, and as one who was privileged to publish in the first volume of the *Alberta Journal of Educational Research*, to comment briefly about the Faculty of Education in this issue commemorating the 50th Anniversary of our Faculty.

First, let me provide you with a bit of information about our history. The Faculty of Education at the University of Alberta was established on April 11, 1942. Therefore, during the academic term 1991-1992, the Faculty and its friends are celebrating the 50th Anniversary of an event that is significant not only to education in the Province of Alberta, but also to the development of education in Canada and in countries around the world.

Our Faculty had modest beginnings. The first course in education at the University of Alberta was approved by the Senate in 1911. By 1922 instruction was being offered in the philosophy of education, educational psychology, history of education, and educational administration. These early efforts evolved in stages through a School of Education administered by the Faculty of Arts and Science, a College of Education, and finally, in 1942, the Faculty of Education.

Creation of the Faculty set the stage for the decision by the province to close its normal schools in Calgary and Edmonton and to turn control of teacher education over to the University of Alberta in 1945. Alberta was the first province in Canada to make teacher education exclusively a university responsibility. The Faculty became the largest and strongest in the country and gained strong governmental and university support not enjoyed by most such faculties in Canada.

While I am in this boasting mode, let me tell you a few more facts about the Faculty of Education at the University of Alberta:

- during its 50-year history, it has awarded over 39,000 BEd degrees;
- approximately one third of the deans of faculties of education in Canadian universities are graduates of this Faculty;
- the Faculty has supplied more faculty members for faculties of education across Canada than any other institution;
- this year we have in attendance approximately 5,000 students, 5,000 undergraduates, and 1,000 graduates (700 MEd and 300 PhD);
- the teaching staff has grown from three full-time instructors in 1942 to an FTE well over 200 in the seven teaching departments;
- during 1990-1991, the Faculty of Education made 4,659 practicum placements;
- in 1942 the Faculty offered 12 education courses; in 1992 the Faculty offers 332 courses;

- in 1989-1990 graduate degrees in education were awarded to residents of 17 countries: Australia, Belize, Canada, Ghana, Hungary, Iceland, Jamaica, Japan, Mexico, Papua New Guinea, The People's Republic of China, South Korea, Sri Lanka, Tanzania, Thailand, the United States of America, and Zimbabwe;
- the Faculty has been involved in teacher preparation programs in Thailand, Tanzania, Korea, China, Uganda, Nepal, India, and Australia.

In closing let me compliment the present editor Dr. Rob Short and all his predecessors for the important role they have played in creating and sustaining a journal of such wide acclaim. Academic communities are concerned with the dissemination of knowledge as well as the extension of knowledge. Teaching is our main concern, but research (the application of systematic intellectual curiosity applied to the borders of knowledge) is required by the teaching function. I am convinced that the *AJER* has contributed significantly to both of these roles. These contributions over the life of the Journal since its inception in 1955 have been recognized and noted by many world class scholars. Thus I consider it appropriate to applaud the *AJER* by citing the oft quoted expression, "You've come a long way, baby!" Keep up the excellent job!

Harvey Zingle

Editor's Note

I would like to take this opportunity as the new editor of *AJER* to thank my immediate predecessor Tom Maguire for the enormous effort, skill, and enthusiasm that he brought to the position. Taking over from Tom is a daunting task. Everything that he accomplished in his three years as editor was with a thoroughness and ability impossible to equal. This issue of *AJER* was planned under Tom's stewardship. To commemorate the 50th Anniversary of the Faculty of Education at the University of Alberta he asked six members of the Faculty who are actively involved in educational research to write articles for the Journal. I believe that the resulting contributions reflect the wide range of research activity and interests at the University of Alberta. Thank you, Tom.

RHS

The Alberta Journal of Educational Research *Vol. XXXVII, No. 4, December 1991, 297-321*

ERWIN MIKLOS
University of Alberta

Trends in Doctoral Research in Educational Administration

A review of 319 doctoral theses in educational administration completed at the University of Alberta between 1958 and 1990 revealed that both the research topics and methodologies were highly varied. The two main goals of the research were to describe and to study relationships. Two main research strategies—survey and case study—were dominant. Questionnaires, interviews, and documents were widely used as sources of data, frequently in combination. Over a period of more than three decades, the proportion of descriptive studies increased while the proportion of relational studies declined. The proportion of studies using case study strategies, interviews, and multiple methods of data collection also increased. When considered in combination, the trends indicate major shifts in the nature of research conducted by doctoral candidates in educational administration.

Educational administration as an area of specialized study has a relatively short history. In most Canadian universities, graduate programs in educational administration were established only in the 1960s and 1970s. Although research is an essential aspect of graduate study, the field has been slow to focus attention on the nature of those research activities. Until recently, there were virtually no books specifically on the subject of research in educational administration. While there have been reviews of research in various areas, critiques of the nature of the research are at times based more on general impressions than on concrete data. Considerations such as these prompted a project that promised to offer additional insights into the state of research in the field of educational administration.[1]

Purpose and Approach

The general purpose of the project on which this paper is based was to review a series of doctoral theses in order to gain a better understanding of the nature of research actually conducted in and on educational administration. The review was restricted to doctoral theses completed in the Department of Educational Administration at the University of Alberta from the time of submission of the first thesis in 1958 to the end of 1990. In some respects, the project may be viewed as a form of case study of research in educational administration.

Among the specific questions defined to guide the review were the following: What are the general characteristics of doctoral research in educational

Erwin Miklos was formerly a professor in the Department of Educational Administration. He currently holds a part-time, postretirement appointment that involves him in both teaching and research activities in the Department.

administration? What purposes, strategies, and methodologies characterize the research? What trends, if any, are evident over time in the characteristics of the research? What insights does the doctoral research offer into the general nature and evolution of research in educational administration? Additional research questions focused on the substantive themes and outcomes of the studies.

The approach that was adopted to addressing the research questions involved identifying a number of aspects of each thesis including the problem statement, the conceptual basis of the study, the research methodology used, and the major findings. In each of these areas, categories were established based on the characteristics of the studies themselves. Although the analysis was influenced to some degree by a priori categories, the general approach was one of trying to organize the data in a manner that would provide a meaningful representation of the nature of the research. The focus of this article is on the characteristics rather than on the substance of the studies. In order to identify trends, results of the analysis are presented for three time periods—1958 to 1970, 1971 to 1980, and 1981 to 1990—that are referred to on occasion as *decades*.

Following a comprehensive analysis of the characteristics of the research, the article shifts to a discussion of research emphasis and trends before ending with a concluding note.

Characteristics of the Research

The research reported in doctoral theses is characterized by diversity. A multiple classification scheme was required in order to develop an adequate description of the nature of the research. The basis of the description selected for this analysis revolved around the purposes served by the research, the general research strategy, and the essential features of approaches to the collection and analysis of data. Although terminology adopted for these descriptions was intended to be consistent with that in general use, some slight departures were necessary in order to reflect accurately the distinctive aspects of the research.

Types of Studies Based on Purpose

Research projects differ in terms of general goals; consequently, one basis on which studies may be classified relates to the purposes identified by the researcher. A preliminary examination of the theses under review suggested that six purposes would provide an adequate basis for a classification. The terms selected for the typology—descriptive, relational, evaluative, developmental, conceptual, and methodological—reflect the major possibilities in the intended outcomes of a particular research project. While most studies could be accurately described by single terms, combinations of terms such as descriptive/relational were required in order to categorize studies with multiple objectives. The placement of theses in specific categories was complicated in some instances by lack of clarity about the intent of the study. Consequently, the substance of the entire thesis and not just the statement of purpose was assessed in determining the category in which a particular study might accurately be placed.

Descriptive studies. In the studies categorized as descriptive, the intent of the researcher was to describe a particular event, practice, process, group, organization, or other phenomenon. Many descriptive studies might more appropriately have been labeled descriptive/analytical because researchers tended to go beyond description to probe for similarities or differences, to identify trends, and to offer possible explanations for findings. Studies in the descriptive category covered a broad range of research into a variety of specific topics.

The variety both in areas of interest and in more specific objectives of doctoral candidates for doing descriptive research is indicated by some illustrative themes and examples. Descriptive research has been conducted

- to determine the extent to which certain developments or changes in education have taken place, such as those on the regionalization of educational administration in various provinces (Chapman, 1974) or on community schools (Sullivan, 1976);
- to trace historically a particular development such as provisions for financing education (Warren, 1962), evolution of the formal structure of separate schools (Fenske, 1968), or development of native education (Daniels, 1973);
- to describe the process of how a decision was made or how a policy was developed and implemented as in Housego's (1964) study of a teacher training issue or Tymko's (1979) analysis of the policy on high school accreditation;
- to describe and analyze practices which are not well understood, for example, control over decisions (March, 1981), behavior of administrators (Duignan, 1979), leadership (Thorpe, 1989), or the selection of superintendents (Van der Linde, 1989);
- to determine attitudes towards or beliefs about various aspects of the operation of educational systems such as financial constraint (Ell, 1988) or inservice education (Pansegrau, 1983);
- to develop guidelines for practice based on a consolidation and interpretation of data from various sources such as legal studies that involves an examination of statutory law and court cases to identify basic principles that have implications for administrative practice (Bargen, 1959; Cooze, 1989; Enns, 1961); and,
- to describe specific organizations (Valentine, 1988) or organizational processes such as change (Alpern, 1986).

Many of the theses that were placed in the descriptive category were justified in terms of contributing to knowledge through providing a description of practice as a point of departure for identifying possibilities for change and improvement.

Relational studies. The primary purpose of the researchers who completed relational studies was, of course, to determine the nature of the relationships between or among certain variables. In general, all of the researchers whose theses fell into this category attempted to establish the plausibility of the relationship to be tested; however, many tended to be cautious and described their studies as "exploratory" in nature. While some

stated hypotheses explicitly, others left them implicit; questions rather than hypotheses were frequently used to guide the data analysis.

As is indicated by the examples, relational studies have focused on a diversity of areas including

- interrelationships among characteristics or behaviors of groups and individuals such as professional role orientation and social interaction (Scharf, 1967);
- interrelationships among characteristics of administrators such as personality variables and innovativeness (Marion, 1966), and principals' job satisfaction and perceptions of effectiveness (Johnson, 1988);
- relationship of organizational variables to individual variables, for example, school climate and teacher behavior (Harvey, 1965) or values and organizational climate (Lupini, 1965);
- relationship of structural characteristics to some measures or indicators of effectiveness (Rourke, 1983); and,
- relationships involving environmental variables such as demographic factors and school attendance (Drolet, 1961).

Although the studies cited above are primarily relational, some also have description as an implicit purpose.

Descriptive/relational studies. A significant proportion of doctoral thesis research might best be described as descriptive/relational in that the researcher attempted to achieve two purposes: to present descriptive data on an area of interest and to examine relationships among selected variables. Researchers who have completed descriptive/relational studies defined the specific purpose in a variety of ways. Some did include *to describe* as an explicit part of the problem statement. However, others tended to use a more general term such as *to analyze, to examine, to explore, to investigate, to identify,* or *to compare.*

As is true of both descriptive and relational studies, descriptive/relational research problems have addressed various specific topics. Among these has been research conducted

- to analyze structure and performance of school organizations (Hersom, 1969);
- to examine the nature of member involvement in a teachers' organization (Ingram, 1965) or burnout as a stress response among principals (MacPherson, 1985);
- to investigate relationships between professional role orientations and bureaucratic characteristics (Robinson, 1966) or job satisfaction of school principals (Rice, 1978); and,
- to identify and describe the nature of prestige hierarchies within the teaching profession (Bride, 1973).

Other examples would be similar to those identified in the discussion of relational studies.

Evaluative studies. In this classification, evaluative studies were those that revolved around assessing the merits of a particular program or practice. The purpose or intent of the researcher was to examine and to make judgments about consequences, outcomes, or effects. Some evaluations were

directed specifically at innovations or new programs while others focused on continuing practices.

The conceptual bases of evaluative studies may reside only partially in a particular theory of evaluation and may draw on other areas as well. Dickie (1983) based his evaluative study on a particular evaluation model, whereas Fraser (1974) and Bethel (1981) based their research on the more general evaluation literature as well as that on organizational effectiveness. Still others, such as Aidoo-Taylor (1986), developed a model for assessing the practice under investigation.

Some evaluative studies had a sufficiently strong descriptive component to warrant the designation of descriptive/evaluative studies. They had no distinctive characteristics other than this particular set of dual objectives.

Developmental studies. In general, developmental projects had the characteristics of what is usually referred to as research and development in education. Developmental projects were oriented primarily toward designing a "product"—a procedure, a program, or a policy—assumed to be of direct interest to the practice of administration. Some testing or evaluation of the product was usually involved.

An indication of the range of possibilities is provided by some of the studies that fell into this category:

- a computer-based model of supply of and demand for teachers (Richards, 1971);
- a program information system for a college (Murphy, 1976), a model for monitoring an educational system (Collett, 1981), and a computer support system for an instructional department (Montgomerie, 1981);
- a program in moral reasoning for use in administrator preparation (Sinclair, 1978); and,
- policy recommendations for the professional preparation of educational administrators (Duncan, 1986).

The conceptual basis of such projects usually was drawn from the relevant literature. In designing the product, the researchers moved from the conceptual to the applied stage.

Conceptual studies. All research has a conceptual base, and some conceptual implications usually can be derived from the results of any particular study. In this regard, all doctoral theses are "conceptual" in some respects. However, certain studies were designed primarily for the purpose of contributing to conceptual development or clarification, and the category of conceptual studies was restricted to those specifically dedicated to achieving this purpose. In a conceptual study, the intent of the researcher was to test or to analyze critically an existing theory or to develop a new conceptualization, construct, or theory.

The general nature of conceptual studies is illustrated by the following examples:

- testing the applicability or validity of a particular theory or model such as the bureaucratic model (Allison, 1980; MacKay, 1964), Fiedler's theory of leadership (McNamara, 1968), or selected models of change (Peters, 1986);

- clarifying concepts of faculty vitality (Nicolson, 1988) or of different approaches to planning (Moore, 1973);
- generating hypotheses or theories related to areas such as the developmental stages of teachers' growth during their careers (Williams, 1986), or the process of educational innovation (Prebble, 1975); and,
- developing an alternative conceptual model such as of organization (Deblois, 1976) or of the policy making process (Kunjbehari, 1981).

Studies that had a conceptual purpose were usually designed in combination with other objectives such as describing certain phenomena, exploring relationships, or developing methodologies.

Methodological studies. In a methodological study, the intent of the researcher was primarily to test the applicability of a research procedure or to develop a new research technique. Although various theses may have made a contribution to the development of research methods, only those that had a fairly explicit methodological objective were placed in this category.

Methodological studies have been directed specifically toward developing instruments, testing research procedures, or developing a new procedure. The specific objectives of methodological studies have taken forms such as

- to construct measures of the organizational structure and operation of secondary schools (Kelsey, 1973), or to develop an instrument to describe dimensions of the teaching-learning process (Stryde, 1973);
- to test the applicability of the techniques developed by Bellack for the analysis of verbal interaction at school board meetings (Holdaway, 1968); and,
- to develop a procedure for examining structural changes in organizations over time (Heron, 1972), or a classification system and a method of analyzing the content of communications (Adams, 1973).

Most studies that placed a significant emphasis on the methodological purpose also gave some attention to the conceptual aspect.

Trends in types of studies. The distributions by percentages of the 319 theses for each type of study across three time periods—1958 to 1970, 1971 to 1980, and 1981 to 1990—are shown in Table 1. About 46% of all theses are primarily descriptive in nature, 16% are relational, and 23% are descriptive/relational. The other types—evaluative, developmental, conceptual, methodological—both singly and in combination—comprise only about 15% of all theses. Only 6% of theses have an evaluative component, and only about 8% have either a conceptual or a methodological focus. Only 2% were classified as being developmental. Although varied intents were evident in the research pursued by doctoral candidates, the main purposes were to describe and to study relationships.

An examination of the percentages across the three decades reveals some changes in types over time. The proportions of descriptive studies increased from about one third in the first decade to well over half in the third, whereas relational studies declined from over one in four to about one in 12. The proportion of descriptive/relational studies decreased from the first to the second decade and then remained steady. The shift in types would seem to be from a decreasing emphasis on relational studies to an increasing emphasis

Table 1
Percentage Distribution of Theses by Type Based on Purpose

Type of Study	1958-1970 N=74	1971-1980 N=112	1981-1990 N=133	Total N=319
Descriptive	32.4%	42.0%	56.4%	45.8%
Relational	27.0	17.0	8.3	15.7
Descriptive/Relational	31.1	19.6	21.1	22.9
Evaluative	0	4.5	6.0	4.1
Descriptive/Evaluative	1.4	1.8	2.3	1.9
Developmental	0	2.7	2.3	1.9
Conceptual	0	2.7	0.8	1.3
Conceptual/Descriptive	0	2.7	3.0	2.2
Conceptual/Descriptive/Methodological	0	2.7	0	0.9
Conceptual/Methodological/Relational	1.4	1.8	0	0.9
Conceptual/Relational	2.7	2.7	0	1.6
Methodological/Relational	4.1	0	0	0.9

on descriptive studies. Evaluative studies emerged primarily in the second decade but did not increase substantially in the third. No patterns are evident in the other types of studies because of the low proportions.

Research Strategies

A second basis for classifying theses relates to the general strategy, method, or design that characterized the research project. Seven major research strategies provided a comprehensive basis for categorizing the thesis research: survey, case, case/survey, comparative, causal-comparative, experimental and panel. Only 10 of the 319 theses could not be placed into one of these categories necessitating an "other" category.

Surveys. Researchers who used a survey strategy collected data from a sample in order to draw conclusions about the population from which the sample was selected. A major variation in survey research related to the sampling procedures. Variety in approaches to sampling rather than a convergence toward a limited number of methods was the general characteristic of survey strategies in thesis research.

Random sampling was used in surveys that related to relatively large as well as to smaller populations. Some studies involved a random sample of teachers in a province, whereas others were somewhat more limited as a random sample of all principals in a province, or still more limited such as a sample of academic staff in one university. Stratified sampling was used to select samples of administrators, samples of regional offices of education, and samples of teachers and administrators. All members of a population were the focus of some surveys (March, 1981; Van der Linde, 1989).

Cluster sampling has been used as a manageable and defensible strategy for collecting data on variables of interest to researchers in educational

administration. Sometimes there may be sequential cluster sampling; for example, Greenfield (1963) selected a sample of 22 districts, two schools in each district, and two classrooms in each school and then focused on the students in those classrooms. Purposive sampling has also been used. For example, Larsen (1984) selected a sample of 10 nurses in a specified age-range who held earned doctoral degrees. Among other approaches to sampling that were used by researchers are quota sampling, reputational sampling, volunteers, and snowball sampling.

Some surveys involved relatively large samples from a particular school district or a particular organization. For example, both Williams (1981) and Sarros (1986) conducted surveys in one school district involving fairly large samples. In contrast, some surveys have involved only a relatively small number of respondents.

Case studies. Research projects that involved the in-depth examination of one particular individual, group of individuals, organization, event, or series of events were categorized as case studies. A source of variation in case studies was the extent to which a historical perspective entered into the analysis. Some case studies were essentially descriptions of a particular event or phenomenon at one point in time. Examples of this type of case study included the analysis of interpersonal influence in a school (House, 1966), the operation of an urban school board (Martin, 1968), the adoption of a participative governance approach in a college (Kelly, 1973), and the operation of community schools (Quarshie, 1989).

In another type of case study, the historical dimension entered to a degree in that there was an attempt to describe developments over a period of time. Most of the studies on how a decision was made or on how a certain policy was developed or implemented are historical in the sense that some efforts were usually made to describe the antecedents of the events that were of particular interest as well as to trace the evolution of the development. For example, Korteweg (1972) focused on a "decade of social studies curriculum development in Alberta," Kozakewich (1980) linked policy developments to activities of interest groups in two previous decades, and Iles (1984) examined the history of school jurisdictional issues in a particular area to identify conflictual themes.

The third type of case study had a more explicit historical focus in that the researcher sought to describe and search for explanations of what occurred and how this came about. For example, Toombs (1966) analyzed parliamentary debates on federal financial participation for the period 1867-1960, Hodgson (1964) examined the nature and purposes of public schools in the Northwest Territories and Alberta from 1885 to 1963, and Scott (1973) focused on the development of secondary education in Jamaica over a 10-year period.

In pursuing research topics that lent themselves to a case study strategy, some researchers chose to include more than one instance or case in the investigation. These are, in effect, multi-site or multi-case studies. For example, Thorpe (1989) included seven administrators in a study of leadership while Young (1989) developed a biographical study around the life experiences of four female educators.

Case/surveys. Some studies had certain characteristics of a case study in that the focus was on one organization; however, data collection procedures resembled those of a survey. Consequently, there was need for a case/survey category to describe this type of study. Some examples may serve to clarify this research strategy.

Eastcott (1975) used a case/survey approach to examine the perceptions of and preferences for participation in decision making of the members of a university faculty. Since only one particular organizational unit was involved, the study might have been developed along the lines of a case study. However, the researcher was guided by questions and used data collection techniques frequently associated with surveys. A study by Jeffrey (1989) of the relationships between the Alberta Teachers' Association and other organizations also exemplifies the combination of case and survey research strategies.

Comparative studies. Even though comparison was an aspect of data analysis in numerous studies, the category of comparative studies was restricted to those that involved a deliberate strategy of making comparisons between two or more sources of data: individuals, organizations, events, or other phenomena. The decision on placing theses into this category was complicated by the variety of ways in which comparisons entered into the design of studies. Some researchers described their studies as comparative when this was more descriptive of the approach to data analysis rather than to overall research strategy. In others, the comparative nature of the design was not dealt with explicitly by the researcher.

Among those researchers who made the comparative nature of the study explicit was Maddocks (1972) in an analysis of approaches to planning development in four postsecondary education systems. Other studies that fell into the comparative category included Sadighian's (1975) study of university goals and governance in North America and Iran, and Mainali's (1985) study of organizational structures in two different cultural contexts.

Causal-comparative studies. Ex post facto or causal-comparative studies involved the use of *naturally-occurring experiments* in which groups were selected because they differed in some characteristic, and an attempt was then made to identify factors that might account for the difference. A relatively small number of studies have been conducted that clearly fell into this category.

In a study of the relationships among leader behavior, staff morale, and productivity, Keeler (1961) assigned a productivity score to each of 179 schools based on the results of grade 9 examinations from which the effects of pupil intelligence test scores had been partialled out. He then selected 23 schools that were above the mean on productivity and 23 schools below the mean as the sample for examining the relationships that were of interest. A similar strategy was used by Bevan (1970) in selecting two high schools that were rated at the extremes among nine schools on openness of climate in a study of the need for independence in high school students. A third example of a causal-comparative strategy is the study by Sullivan (1987) of variables related to high school achievement. The researcher selected samples of 48 graduates and 42 dropouts in one school district and then explored differen-

ces between the two groups on students' home environment, personal characteristics, academic performance, and school life experiences.

Experimental studies. In an experimental study the researcher engaged in the intentional and controlled manipulation of variables to test for a cause-effect relationship. A number of theses have been completed that have the characteristics of experimental studies; others are a closer approximation of quasi-experimental in that the random selection to treatment groups was absent.

In the first experimental study conducted by a doctoral candidate, Brown (1961) examined the differential effect of stress-inducing supervision on student teachers. Berghofer (1972) used three groups of graduate students to test the effects of different levels of exposure to literature in the area of futures research on the development of educational policy. Graduate students were also used by Plaxton (1969) to study the effects of different decision rules on interaction in small groups. Batchler (1977) involved workshop participants in a simulation to study the relationship between organizational climate and administrative behavior.

Panel studies. Panel strategies involve collecting data from the same set of individuals at different points in time. The closest approximation to panel studies in the theses reviewed were Delphi studies. In these studies, participants who are expert or knowledgeable in an area were asked to assess the probability, feasibility, or desirability of future trends or developments. Members of the panel then participated in successive rounds of data collection in which results from the preceding round were made available to the respondents.

The first task of the researcher was to identify prospective panel members on the basis of some criteria. Barrington (1980) used a panel of 10 college presidents and seven government officials to assess the projected impact of environmental forces on community colleges. Other researchers have used the same basic strategy to identify appropriate respondents to study policy issues in specific areas.

Trends in research strategies. The percentage distributions of theses in Table 2 indicates that two major strategies might merit the designation of traditions: the survey and the case study. Together these strategies account for nearly 85% of all doctoral research, and when the case/survey strategies are included the proportion approaches 88%. Although making comparisons was part of the analysis in a number of studies, only about 3% placed a major emphasis on a comparative design. Together experimental and causal-comparative research strategies were used in only 10 theses, while five involved data collection from a panel. The "other" category included simulations, computer modeling, information systems, and semiotic analysis.

Trends over the three decades reveal a decline in the proportion of surveys and an increase in the proportion of case studies. Surveys accounted for approximately 70% of studies completed in the first decade and just under 50% of those completed in the second and third decades. Case studies increased from around one fifth in the first decade to more than one third in the most recent decade.

Table 2
Percentage Distribution of Theses by Research Strategy

Research Strategy	1958-1970 N=74	1971-1980 N=112	1981-1990 N=133	Total N=319
Survey	71.6%	48.2%	48.9%	53.9%
Case	20.3	31.2	35.3	30.4
Case/Survey	1.4	5.4	4.5	4.1
Comparative	0	5.4	3.8	3.4
Experimental	2.7	4.5	0	2.2
Panel	0	0.9	3.8	1.9
Causal-Comparative	2.7	0	0.8	0.9
Other	1.4	4.5	3.0	3.1

Analysis not reported in the table revealed that case studies that did not have an explicit historical dimension increased across the three decades. Studies that partially had a historical dimension decreased between the first and second decade and recovered only slightly in the third decade. In contrast, those that had a strong and explicit historical dimension increased from the first to the second decade and decreased slightly in the third.

Data Collection Methods

Although research methods could be described broadly in terms of such categories as quantitative and qualitative, a more precise description can be obtained from a consideration of the major methods of data collection or sources of data. For the types of studies carried out by doctoral candidates, data collection involved five main methods or sources that were used either singly or in combination: questionnaires, interviews, documents, observations, and records.

Questionnaires. As used here, the term questionnaire includes both various published or standardized instruments as well as those designed by the researcher. The number of researchers who designed their own questionnaires was considerably greater than those who modified available instruments or used published instruments.

The published instruments that have been used include a range of tests or scales of a primarily attitudinal nature. Among these are the Rokeach Dogmatism Scale, Edwards Personal Preference Schedule, Differential Values Inventory, the Kluckhohn value scale, and the Maslach Burnout Inventory. Researchers have also used instruments more closely associated with studies in administration such as the Leader Behavior Description Questionnaire, the Organizational Climate Description Questionnaire, and the measures associated with Fiedler's leadership theory. Although some of these instruments have tended to be overused in the field generally, this was not the case in relation to this particular set of studies.

Instruments that were adapted to the educational setting resulted in the development of a number of lines of research. For example, MacKay's (1964) adaptation of Hall's Organizational Inventory was used later by other researchers such as Robinson (1966), Kolesar (1967) and Richert (1968). The adaptation of Corwin's Professional Role Orientation scale developed by Hrynyk (1966) was also used subsequently by a number of researchers (Masse, 1969; Nixon, 1975; Scharf, 1967). Such lines of research have seldom resulted in the systematic and cumulative analysis of the merits of the instruments involved.

Interviews. The interview was used in a variety of ways ranging from being the primary source of data to being supplementary to other methods of data collection. Regardless of their general place in the research design, interviews varied in degree of structure from relatively highly structured to unstructured. Among the more highly structured forms of interviews were those associated with the Aston methodology for developing measures of the structure of organizations which was used by Newberry (1971), Heron (1972), Kelsey (1973), Sackney (1976) and Rourke (1983). Other researchers described their interviews as either structured, semistructured or relatively unstructured. The majority of interviews fell into the semistructured area in that the researcher usually had an interview guide or schedule but remained flexible, allowing the questions to shift in sequence. Relatively few researchers regarded their approach as being highly unstructured.

Observations. As a method of data collection, observation has taken a variety of forms. Like interviews, observations included both structured observation and those in which there was no predetermined basis for categorizing the observations.

Both direct and indirect forms of observation were used in structured approaches. As examples of direct observation, Mathews (1967) personally observed school board meetings and used interaction analysis to code the process. Similarly, Hemphill (1968) used the same coding methods but involved two trained coders to collect data on school board meetings. In contrast, Plaxton (1969) videotaped group sessions and then analyzed the interaction. Both direct observations and analyses of videotapes were used in studies of teaching and learning processes (King, 1979; Mireau, 1980; Wodlinger, 1980). Some researchers (Allan, 1985; Cooper, 1986; Duignan, 1979) used structured observation techniques—coding within predetermined categories—to analyze the work of administrators. In contrast to this approach, Moylan (1988) used more varied methods, as did Bird (1982) and Hall (1989) to study particular individuals.

In using observational strategies, researchers have defined varying participant roles for themselves in order to carry out studies in a variety of sites. A number of researchers have used observation in the form of attendance at meetings or becoming involved in other activities. Others were actual participants in the events that were the focus of their research. In such instances observation may have been a supplementary rather than a major aspect of data collection.

Documents. Documents as data sources included a broad range of written or printed materials that were relevant to a particular subject. Examples of

the types of documents that have served as primary and secondary sources are listed below:

- official files, reports, transcripts of proceedings, records of parliamentary debates, minutes, archives;
- legal documents such as laws and by-laws, acts of governments, law reports, statutes, records of court cases;
- correspondence, letters, briefs, memoranda;
- newspaper articles, newsletters, bulletins, periodicals ;
- policy documents, reports, historical documents; and,
- literature on a particular theme, major writings of theorists of interest.

Documentary sources were subjected to an appropriate form of content analysis to extract the data relevant to the research problem.

Records. For purposes of this classification, *records* was defined in the narrow sense to refer to data sources that are primarily quantitative in nature and that have an official or authoritative character. The following are among the types of records that have been used as data sources:

- student achievement data, student information files, cumulative records;
- statistical records such as earnings of various job classifications, census data, and vital statistics;
- annual reports of government departments, operational reports submitted by school districts; and,
- audited financial statements.

Data extracted from records were analyzed by a variety of descriptive techniques or involved inferential analysis.

Trends in use of methods/sources. The extent to which questionnaires, interviews, observations, documents, and records were used in data collection across the 319 theses is indicated by the percentage distributions in Table 3.

Nearly 60% of the research projects involved data collection by questionnaires followed closely by interviews that were used in about 56% of the theses and by documents that were data sources in about 38% of the studies.

Table 3
Percentage of Use of Data Collection Methods or Sources

Method or Source	1958-1970 N=74	1971-1980 N=112	1981-1990 N=133	Total N=319
Questionnaires	67.6%	59.8%	53.4%	58.9%
Interviews	27.0	54.5	72.9	55.8
Documents	31.1	37.5	42.9	38.2
Observations	4.1	15.2	19.5	14.4
Records	20.3	9.8	9.8	12.2

Note: Sums of columns exceed 100% because of use of multiple methods.

Both observations and records were used in relatively small proportions of the theses, about 14% and 12% respectively.

A comparison of the proportion of research projects that used particular methods indicates changes across the three decades. The extent to which questionnaires were involved in data collection declined from being used in more than two thirds of theses in the first decade to just over one half in the third decade. Although relatively small, the proportion of studies involving use of records also declined. In contrast, the use of interviews increased from slightly more than one fourth of theses to nearly three fourths. Observations also increased dramatically as sources of data, while the use of documents increased from just over 30% of theses to about 43%. In general, the trends indicate a fairly dramatic shift from quantitative to qualitative methods across the three decades.

The extent to which multiple methods of data collection were used is indicated by the data in Table 4. About 42% of the studies used a single method of data collection while about 38% used two methods. Nearly 19% used three methods, and only a small proportion of research projects involved more than three methods of data collection. Over the three decades, the proportion of studies using only a single method of data collection was halved, from about 63% to 31%. In contrast the number of studies involving two methods almost doubled from 23% to 42%. Although fewer studies used three methods, the increase was proportionately about the same.

The particular method or combination of methods used by the 319 researchers is indicated in Table 5. Eighty-five of the 134 studies that used only a single method involved questionnaires, while documents and records were the single data source in each of 20 studies. Eight studies used interviews exclusively, and only one used observation exclusively. The decline in the use of questionnaires over the three decades noted previously is also evident in this table.

The two most frequently occurring combination of methods were interviews/questionnaires and documents/interviews, each of which comprised about 15% of the studies. The next most frequently occurring combinations were documents/interviews/questionnaires and documents/interviews/observation representing about 8% and about 6% of the studies respectively. Various other combinations occurred but in total represent only about 14% of all studies.

Table 4
Percentage Distribution of Theses by Number of Methods

Number of Methods	1958-1970 N=74	1971-1980 N=112	1981-1990 N=133	Total N=319
One	63.5%	41.1%	30.8%	42.0%
Two	23.0	42.9	42.1	37.9
Three	13.5	14.3	24.8	18.5
Four	0	1.8	2.3	1.6

Table 5
Distribution of Theses by Data Collection Method

Method	1958-1970	1971-1980	1981-1990	Total
Questionnaires	32	29	24	85
Documents	7	8	5	20
Records	8	7	5	20
Interviews	0	1	7	8
Observations	0	1	0	1
Interview/Questionnaire	4	18	27	49
Document/Interview	6	20	22	48
Document/Interview/Questionnaire	7	8	11	26
Document/Interview/Observation	0	3	15	18
Interview/Observation	1	4	4	9
Interview/Questionnaire/Records	1	3	3	7
Questionnaire/Records	4	1	1	6
Observation/Questionnaire	1	4	0	5
Document/Interview/Observation/Quest.	0	2	3	5
Interview/Observation/Questionnaire	0	2	1	3
Interview/Observation/Records	0	0	3	3
Documents/Observation	0	1	0	1
Documents/Records	1	0	0	1
Interview/Records	0	0	1	1
Documents/Interviews/Records	1	0	0	1
Document//Observation/Questionnaire	1	0	0	1
Documents/Questionnaire	0	0	1	1
Total	74	112	133	319

Purpose, Strategy, and Method

Further insights into the characteristics of research in general and trends over time may be obtained by combining the different bases of classification. This was done by focusing on purpose and research strategy. To simplify the presentation, the description is restricted to major categories or types of studies.

When a cross classification of purpose and strategy was developed, four major types were evident: descriptive case studies, descriptive surveys, descriptive/relational surveys, and relational surveys. Nearly three fourths of the 319 theses could be placed into one of these four types. Approximately 26% of all theses were descriptive case studies and another 18% were descriptive/relational surveys. Descriptive surveys and relational surveys accounted for about 16% and 12% respectively of the studies.

Table 6
Percentage Distribution of Theses by Type Based on Purpose/Strategy

Type	1958-1970 N=74	1971-1980 N=112	1981-1990 N=133	Total N=319
Descriptive Case	20.3%	24.1%	30.8%	26.0%
Descriptive/Relational Survey	29.7	15.2	15.0	18.5
Descriptive Survey	12.2	13.4	19.5	15.7
Relational Survey	21.6	12.5	7.5	12.5
Other	16.2	34.8	27.1	27.3

Descriptive case studies increased in each decade from about 20% through 24% to 31% as proportions of all studies. A similar trend—12% to 13% and then to 20%—is evident for descriptive surveys. In contrast, descriptive/relational surveys declined from about 30% to 15%. Even more dramatic was the decline of relational surveys from 22% to less than 8% of all studies. Of interest also is the high proportion of studies, nearly 35%, in the "other" category during the second decade.

Research Topics

Because there is no consensus about the boundaries or major subdivisions of the field, any basis for categorizing theses in terms of the substantive focus is likely to appear to be somewhat arbitrary. The categories and subcategories for organizing the substantive review emerged as the result of a lengthy process of sorting, classifying, and regrouping abstracts of the studies. Among the underlying objectives in the procedure was the desire to present as unified a perspective on the research as possible, to be parsimonious in the topics used, and, at the same time, to represent adequately the diversity of the research areas. Inevitably, the major research topics identified reflect a particular view of the subdivisions in the study of educational administration.

Through the process that was adopted, seven major topics were identified: individual and organization, organizational analysis, policy development and implementation, educational and organizational change, the context of educational administration, decision making, and the administrator and administration. As might be anticipated, some arbitrary decisions were necessary in this process. Certain studies could easily be placed in any one of a number of categories, while other topics tested the limits of the framework. To have done justice to the uniqueness of the individual theses would have necessitated a much more extensive list of topics and subtopics.

An indication of the general substantive emphasis in the research is provided by the data in Table 7. The general topic that was labeled *individual and organization* encompassed the largest proportion of studies, approximately one fifth. About one sixth of the studies related to each of two topics: organizational analysis and policy studies. The other topics in order of decreasing emphasis were educational and organizational change, the con-

Table 7
Percentage Distribution of Theses by Topic

Topic	1958-1970 N=74	1971-1980 N=112	1981-1990 N=133	Total N=319
Individual and Organization	14.9%	17.9%	24.1%	19.7%
Organizational Analysis	18.9	18.8	14.3	16.9
Policy Development and Implementation	5.4	21.4	19.5	16.9
Educational and Organizational Change	9.5	14.3	16.5	14.1
Context of Educational Administration	23.0	11.6	7.5	12.5
Decision Making	16.2	10.7	8.3	11.0
Administrators and Administration	12.2	5.4	9.8	8.8

text of education, decision making, and the administrator. The latter category included fewer than one tenth of all studies.

Some changes in substantive emphasis are evident over the three decades. Most dramatic is the consistent decrease in the proportion of studies that relate to the context of education —from more than one fifth of the studies in the first decade to about one twelfth in the most recent decade. A similar trend is evident in research related to decision making. The focus on the administrator decreased from the first to the second decade but moved upward in the third decade. Research in the general area of organizational analysis remained steady over the first two decades but declined slightly as a proportion of studies in the third. The remaining three topics—individual and organization, policy development and implementation, and educational and organizational change—all show increases as a proportion of studies from the first to the third decade. The most dramatic change was in the area of policy development and implementation, which accounted for only about 5% of the studies in the first decade but approximately 20% in the second and third. Research into change and into the individual and the organization also increased steadily, as proportions of total studies, across the three decades.

Discussion

Various aspects of the doctoral research warrant some comment; however, the discussion in this paper is restricted to the emphasis and the trends in the research.

Research Emphasis

Although there is a measure of diversity along various dimensions, doctoral theses in educational administration seem to be dominated by two major models of research. The first model supports studies that focus on relationships among variables, use survey strategies, rely on questionnaire methods of data collection, and subject the data to quantitative analysis. The second model promotes studies that describe a particular area of interest or series of events, involve case study strategies, use documentary and interview data sources, and analyze the data through nonquantitative methods. Some of the

variations in the research reflect a blending of certain features of the two models such as using surveys to generate descriptions or supplementing questionnaire data in relational studies with interview and documentary data. The pervasiveness of these models probably reflects to a considerable degree the constraints under which doctoral research is conducted as well as the state of research in the field of study generally.

Given the nature and state of studies in educational administration, some questions might be raised about the merits of the type of research that has been conducted. An interest in describing phenomena may be appropriate for a relatively new field of study, especially one that is closely associated with an area of professional practice. Similarly, examining relationships among variables may reflect a need to conform to a particular conception of "scientific" research in order to establish and maintain academic credibility. However, the limited number of conceptual or methodological studies seems inconsistent with the needs of a developing area of research. Furthermore, the degree of effort that has been directed toward evaluating practices or developing alternatives does not address adequately the needs in the field of practice. If research in educational administration were oriented more toward improving practice, then the evaluation of existing programs or policies would seem to be an appropriate focus for more doctoral studies. A similar statement might be made about developmental research; more attention might be directed toward developing ideas or innovations that have administrative or educational application.

In spite of recent debates about quantitative and qualitative methods, the characteristics of thesis research suggest that researchers have used both methods extensively and that the two methods have been combined in various studies. The acceptability of, and apparent success with, these different methods may reflect both eclecticism and pragmatism. Researchers who have combined both quantitative and qualitative methods may have avoided addressing some of the more fundamental issues that have been the focus of debates. Different methods of data collection probably were combined within the same philosophical perspective. The tendency to use both quantitative and qualitative methods may be indicative both of allowing the problem to determine the methods of data collection, as well as of an interest in developing more comprehensive descriptions.

An examination of the types of studies conducted in each of the three decades suggests a process of change in approaches to doing research. Significant among these changes is a shift away from relational studies to more descriptive studies, from surveys to case studies, from questionnaires to interviews and documents as data sources, and from single to multiple methods of data collection. These specific changes may be indicative of a significant shift in basic conceptions about the nature of research appropriate for a field such as educational administration. A fairly narrow conception of research may be giving way to broader definitions and alternative conceptions. There is increasing variety in research strategies and methods; established approaches are not disappearing entirely but instead are being supplemented by other strategies. In these developments, the emergence of

an interpretive or hermeneutic orientation such as in Evans' (1989) study may prove to be of particular significance.

Research Trends

In recent years, different orientations in research have sometimes been conceptualized in terms of the notion of alternative paradigms. Similarly, changes in orientations over time have been discussed from the perspective of paradigm shifts. Although variations in approaches to doctoral research in educational administration may not warrant the designation of paradigms, the broader debate prompts questions about the nature of underlying differences and trends. A consideration of the variations in the theses that have been reviewed suggests that they might be differentiated on the basis of three broad clusters of characteristics. The clusters of characteristics reflect different emphases in basic purpose, research strategy, methodology, and stance of the researcher as well as in substantive focus. For the purposes of this discussion, these clusters of characteristics will be referred to as *types* within a loose definition of the term. The types are specific to the theses that have been reviewed and are not proposed as being descriptive of the nature of research in the broader context.

The first type of doctoral study (Type A) is based on questions that have considerable practical import. Studies of this type promise outcomes that have practical or social relevance. More specifically, the research problems draw significance from the world of administrative practice. The purposes of the research project are presented in a form that suggests a descriptive or perhaps a relational study. Research strategies are likely to be surveys or case studies with appropriate reliance on either quantitative or qualitative methods of data collection. Implicitly, the researcher adopts the stance of an objective observer or analyst who will provide knowledge that has potential to increase understanding of a problem and possibly to offer some means for resolving the problem. In general, Type A studies reflect the needs of a field of practice and probably are consistent with the characteristics of the earliest orientation to research in educational administration.

A second type of study (Type B) is associated closely with what has come to be known as the the theory movement. Essentially, the research problems investigated have both theoretical and practical significance. If the problem originates in the field of practice, then it is cast in a theoretical framework. Alternatively, the research problem may originate within a theoretical formulation, and the practical aspects may be but one instance of a class of similar problems. The purpose of the research may involve testing a particular line of theorizing or establishing the basis for developing a new conceptualization. Any one of a broad range of research strategies—case studies, surveys, quasi-experimental designs, and others—may be used. Implicitly, the general stance of the researcher is that of a "scientist" who is capable of contributing objective, conceptually based, and empirically verified knowledge to the study of educational administration.

The third type of study (Type C) may be viewed, in part, as a reaction to Type B research. Although based in the social and human sciences, the definition of science is much broader. The essential difference from Type B

research involves developing understandings more from the perspective of the participant than from that of an objective observer. There is some resemblance to Type A research in that the research is descriptive, but the nature of the understanding sought is different. The research strategies are more likely to be case studies or multi-case studies using a variety of field research techniques. Essentially, the data collection methods are qualitative while the methods of analysis tend to be interpretive.

All three types of research are evident in the doctoral studies that were reviewed; however, the boundaries are not as sharply delineated as the definitions would seem to indicate. There is a considerable degree of overlap and blending of the types. Given this overlap, attempts to place individual studies into particular types would lead to distortions. Nevertheless, conceptualizing the research in terms of these three types offers a basis for describing the changes in emphasis over the three decades.

Some of the research initiated in the early years of the Department reflects Type A characteristics. Studies that resemble Type B research gained prominence in the 1960s and 1970s. Type C studies emerged primarily in the 1980s and represent an increasing proportion of doctoral research. Nevertheless, examples of all three types are evident in recent years, indicating the presence of pluralism in research orientations.

Of particular interest is the presence of studies that combine the characteristics of two or all three types. For example, some studies that have most of the characteristics of Type A research may also reflect the conceptual orientation and methods of Type B research even though these may not be essential to achieving the purposes of the study. Similarly, Type C research may show some characteristics that resemble those of Type B research. A specific illustration is the use of established approaches for determining validity and reliability even though these may be inconsistent with the basic assumptions.

The emergence of Type C studies was preceded by some "erosion" in the canons of Type B research. The fairly systematic approaches involving explicit theorizing and hypothesis testing evident in earlier studies subsequently declined. In particular, an increasing number of studies lacked an explicit theoretical base, and hypotheses were replaced by research questions. Similar to the earlier influences of Type B on Type A research, Type C has also had an influence on Type B studies. This has generally taken the form of adding a qualitative dimension—using open-ended questions and supplementing questionnaire data with interviews or document analysis—to strategies that previously were primarily quantitative .

An interpretation in terms of these types of research suggest that an evolution metaphor is appropriate for understanding the development of doctoral research in educational administration. As new clusters of strategies emerge, they serve to modify existing approaches and, consequently, change the overall character of the research without replacing totally the previous strategies, at least not in the short term. Whether emerging emphases will eventually change the fundamental nature of the research remains to be seen.

The discussion of trends in terms of these types suggests the possibility of other types. For example, to date there is no clear evidence in doctoral studies of research based on a critical analysis of the practice of educational administration or of educational policies. However, there are some indications that such an orientation might emerge from an increased concern, as in Maynes' (1990) study, with the substantive aspects of policy issues and not just the processes through which the issues might be resolved.

Conclusion

Research conducted by graduate students in fulfillment of degree requirements constitutes a major part of the total research effort in the study of educational administration. Within the graduate research area, doctoral studies in particular have potential for making a significant contribution to the further development of the field. Although some doctoral research is published, many of the studies do not attain the necessary degree of visibility to ensure that they come to the attention of other researchers. As a consequence, the full potential of doctoral research is seldom realized. A review of the doctoral theses completed in one department tends to confirm the desirability of carrying individual studies through to publication as well as of allocating resources to synthesizing and critiquing studies. Among the possible benefits of such activity would be a sharper focus on important topics, increased conceptual and methodological rigor, and reduced fragmentation of research activity. Clearly, cooperative and collaborative activities across departments in different universities would ensure even greater returns on the investment in graduate students and in graduate student research.

In addition to contributing to knowledge in specific areas, collectively doctoral studies offer a perspective on the general nature of educational administration research. When viewed over time, doctoral theses offer insights into the development of the field of study. The perspective reveals that while there are some "traditions," the field of study is also dynamic. Collectively, researchers appear to be engaging in a search for ways in which the study of educational administration can contribute to improving administrative practices and to linking those practices to the attainment of desired educational goals. Emerging pluralistic and self-critical orientations promise progress in the continuing development of research in educational administration.

Note

1. A grant for the project from the Support for the Advancement of Scholarship Fund awarded by the Scholarship and Research Awards Committee, Faculty of Education, is gratefully acknowledged. This paper is based in part on a preliminary report on the project identified as follows: Miklos, E. (1990). *Evolution in doctoral research in educational administration at the University of Alberta, 1958-1990* (Draft Version). Edmonton: Department of Educational Administration, University of Alberta.

References

Adams, W.A. (1973). *Communication patterns in formal organizations*. Unpublished doctoral dissertation, University of Alberta.

Aidoo-Taylor, N. (1986). *Staff perceptions of the staff development program at the Northern Alberta Institute of Technology (NAIT)*. Unpublished doctoral dissertation, University of Alberta.

Allan, D.J. (1985). *The administrative behavior of the Alberta college president: A descriptive study*. Unpublished doctoral dissertation, University of Alberta.
Allison, D.J. (1980). *An analysis of the congruency between a model of public schools and Max Weber's model of bureaucracy*. Unpublished doctoral dissertation, University of Alberta.
Alpern, M. (1986). *Cultural change in an inner city school*. Unpublished doctoral dissertation, University of Alberta.
Bargen, P.F. (1959). *The legal status of the Canadian public school pupil*. Unpublished doctoral dissertation, University of Alberta.
Barrington, G.V. (1980). *The impact of environmental forces on Alberta community colleges 1980-1990*. Unpublished doctoral dissertation, University of Alberta.
Batchler, M.W. (1977). *A study of relationships between organizational climates and the administrative behavior of school administrators*. Unpublished doctoral dissertation, University of Alberta.
Berghofer, D.E. (1972). *The futures perspective in educational policy development*. Unpublished doctoral dissertation, University of Alberta.
Bethel, K.M. (1981). *Perceived effectiveness of teacher education programs in the Bahamas*. Unpublished doctoral dissertation, University of Alberta.
Bevan, H.G. (1970). *An empirical study of the need for independence in high school students*. Unpublished doctoral dissertation, University of Alberta.
Bird, D.C. (1982). *A practicum associate: An ethnographic account*. Unpublished doctoral dissertation, University of Alberta.
Bride, K.W. (1973). *A study of prestige and attitude differentials among practicing Alberta educators*. Unpublished doctoral dissertation, University of Alberta.
Brown, A.F. (1961). *The differential effect of stress-inducing supervision on classroom teaching behavior*. Unpublished doctoral dissertation, University of Alberta.
Chapman, R.J. (1974). *Regionalization of the administration of public education in Canada*. Unpublished doctoral dissertation, University of Alberta.
Collett, D.J. (1981). *Testing a model for monitoring an educational system*. Unpublished doctoral dissertation, University of Alberta.
Cooper, J.R. (1986). *Principals at work—A descriptive study of principals of Grades I-XII schools in Zones II and III in the Province of Alberta*. Unpublished doctoral dissertation, University of Alberta.
Cooze, J.C. (1989). *Legal liability of Canadian school boards and teachers for school accidents causing physical injuries to students*. Unpublished doctoral dissertation, University of Alberta.
Daniels, E.R. (1973). *The legal context of Indian education in Canada*. Unpublished doctoral dissertation, University of Alberta.
Deblois, C. (1976). *An emerging model of organization based on the literature of liberation*. Unpublished doctoral dissertation, University of Alberta.
Dickie, B.T. (1983). *Portrait of a practicum*. Unpublished doctoral dissertation, University of Alberta.
Drolet, J.-Y. (1961). *A study of the impact of demographic and socio-economic factors on school attendance rates in the Province of Quebec from 1901 to 1951*. Unpublished doctoral dissertation, University of Alberta.
Duignan, P. (1979). *Administrative behavior of school superintendents: A descriptive study*. Unpublished doctoral dissertation, University of Alberta.
Duncan, D.B. (1986). *Policy recommendations regarding educational administration in Alberta*. Unpublished doctoral dissertation, University of Alberta.
Eastcott, L.R. (1975). *Faculty perceptions of and preferences for participation in decision-making*. Unpublished doctoral dissertation, University of Alberta.
Ell, J.F. (1988). *Faculty perceptions of financial constraint in universities*. Unpublished doctoral dissertation, University of Alberta.
Enns, F. (1961). *The legal status of the Canadian public school board*. Unpublished doctoral dissertation, University of Alberta.
Evans, P.R. (1989). *Ministrative insight: Educational administration as pedagogic practice*. Unpublished doctoral dissertation, University of Alberta.
Fenske, M.R. (1968). *The evolution of the formal structure of separate schools in the Prairie Provinces*. Unpublished doctoral dissertation, University of Alberta.
Fraser, I.E. (1974). *The effectiveness of the Learning Assistance Field Project*. Unpublished doctoral dissertation, University of Alberta.

Greenfield, T.B. (1963). *Systems analysis in education: A factor analysis and analysis of variance of pupil achievement*. Unpublished doctoral dissertation, University of Alberta.
Hall, D.E. (1989). *Interpretive and adaptive strategies in college leadership*. Unpublished doctoral dissertation, University of Alberta.
Harvey, R.F.E. (1965). *School organizational climate and teacher classroom behavior*. Unpublished doctoral dissertation, University of Alberta.
Hemphill, H.D. (1968). *The structure of interaction at school board meetings*. Unpublished doctoral dissertation, University of Alberta.
Heron, R.P. (1972). *Growth stages in the development of college structures*. Unpublished doctoral dissertation, University of Alberta.
Hersom, N.L. (1969). *Dimensions of organizational structure and organizational behavior*. Unpublished doctoral dissertation, University of Alberta.
Hodgson, E.D. (1964). *The nature and purposes of the public school in the Northwest Territories (1885-1905) and Alberta (1905-1963)*. Unpublished doctoral dissertation, University of Alberta.
Holdaway, E.A. (1968). *An analysis of the verbal moves in school board meeting interaction*. Unpublished doctoral dissertation, University of Alberta.
House, J.H. (1966). *An analysis of interpersonal influence relations within a school organization*. Unpublished doctoral dissertation, University of Alberta.
Housego, I.E. (1964). *How a decision was made: A study of the teacher training issue in Saskatchewan*. Unpublished doctoral dissertation, University of Alberta.
Hrynyk, N.P. (1966). *Correlates of professional role orientation in teaching*. Unpublished doctoral dissertation, University of Alberta.
Iles, R.L. (1984). *School jurisdictional reorganization: A naturalistic case study of conflict management*. Unpublished doctoral dissertation, University of Alberta.
Ingram, E.J. (1965). *Member involvement in the Alberta Teachers' Association*. Unpublished doctoral dissertation, University of Alberta.
Jeffrey J.H. (1989). *Linkages and effectiveness of relationships between the Alberta Teachers' Association and other organizations*. Unpublished doctoral dissertation, University of Alberta.
Johnson, N.A. (1988). *Perceptions of effectiveness and principals' job satisfaction in elementary schools*. Unpublished doctoral dissertation, University of Alberta.
Keeler, B.T. (1961). *Dimensions of the leader behavior of principals, staff morale and productivity*. Unpublished doctoral dissertation, University of Alberta.
Kelly, G.O. (1973). *A study of participation in college governance*. Unpublished doctoral dissertation, University of Alberta.
Kelsey, J.G.T. (1973). *Conceptualization and instrumentation for the comparative study of secondary school structure and operation*. Unpublished doctoral dissertation, University of Alberta.
King, L.H. (1979). *An attributional analysis of student achievement-related behavior and the expectancy effect*. Unpublished doctoral dissertation, University of Alberta.
Kolesar, H. (1967). *An empirical study of client alienation in the bureaucratic organization*. Unpublished doctoral dissertation, University of Alberta.
Korteweg, L. (1972). *A decade of social studies curriculum development in Alberta*. Unpublished doctoral dissertation, University of Alberta.
Kozakewich, E.J. (1980). *The Cameron Commission, interest groups and policy-making*. Unpublished doctoral dissertation, University of Alberta.
Kunjbehari, L.L. (1981). *Politics and expertise in policymaking: A model and a case study*. Unpublished doctoral dissertation, University of Alberta.
Larsen, J.B. (1984). *A psychosocial study of the career development of nurses with earned doctoral degrees*. Unpublished doctoral dissertation, University of Alberta.
Lupini, D. (1965). *A study of the relation of differential values to social and administrative interactions*. Unpublished doctoral dissertation, University of Alberta.
MacKay, D.A. (1964). *An empirical study of bureaucratic dimensions and their relation to other characteristics of school organizations*. Unpublished doctoral dissertation, University of Alberta.
MacPherson, M.A. (1985). *Job stress and the school principal: A study of burnout and its relationship to personal, situational, and organizational variables*. Unpublished doctoral dissertation, University of Alberta.

Maddocks, G.R. (1972). *A comparative analysis of approaches to planning development in post-secondary education*. Unpublished doctoral dissertation, University of Alberta.
Mainali, M.P. (1985). *Organizational structure and cultural contexts*. Unpublished doctoral dissertation, University of Alberta.
March, M.E. (1981). *Variations in degree of control over educational decisions*. Unpublished doctoral dissertation, University of Alberta.
Marion, G.B. (1966). *A study of selected factors related to the innovativeness of elementary school principals*. Unpublished doctoral dissertation, University of Alberta.
Martin, S.W. (1968). *A political systems analysis of an urban school board*. Unpublished doctoral dissertation, University of Alberta.
Masse, D. (1969). *Teacher participation and professional attitudes*. Unpublished doctoral dissertation, University of Alberta.
Mathews, N.O. (1967). *A study of the decision-making process of two school boards in an Alberta community*. Unpublished doctoral dissertation, University of Alberta.
Maynes, W.G. (1990). *The education of Edmonton's urban poor: A policy perspective*. Unpublished doctoral dissertation, University of Alberta.
McNamara, V.D. (1968). *The principal's personal leadership style, the school staff leadership situation, and school effectiveness*. Unpublished doctoral dissertation, University of Alberta.
Mireau, L.J. (1980). *Teacher expectancy effects and student attributes*. Unpublished doctoral dissertation, University of Alberta.
Montgomerie, T.C. (1981). *Instructional department computer support systems: A strategy and an implementation*. Unpublished doctoral dissertation, University of Alberta.
Moore, B.D. (1973). *Two approaches to educational planning: Freire and Worth*. Unpublished doctoral dissertation, University of Alberta.
Moylan, J. (1988). *Through a looking glass: A female administrator interprets the perspective of a female administrator*. Unpublished doctoral dissertation, University of Alberta.
Murphy, P.J. (1976). *A community college programme information system*. Unpublished doctoral dissertation, University of Alberta.
Newberry, J.F. (1971). *A comparative analysis of the organizational structure of selected post-secondary educational institutions*. Unpublished doctoral dissertation, University of Alberta.
Nicolson, I.J. (1988). *The concept of faculty vitality*. Unpublished doctoral dissertation, University of Alberta.
Nixon, M.T. (1975). *Women administrators and women teachers: A comparative study*. Unpublished doctoral dissertation, University of Alberta.
Pansegrau, M.V. (1983). *Teachers' perspectives on inservice education*. Unpublished doctoral dissertation, University of Alberta.
Peters, F. (1986). *Assessing models of change: A case study analysis*. Unpublished doctoral dissertation, University of Alberta.
Plaxton, R.P. (1969). *The relationships of decision-rule to interaction patterns, satisfaction, and commitment in small groups*. Unpublished doctoral dissertation, University of Alberta.
Prebble, T.K. (1975). *The Jordan Plan: A case study in educational change*. Unpublished doctoral dissertation, University of Alberta.
Quarshie, J.D. (1989). *A comparative analysis of the operation of community schools in inner city, rural, and suburban settings in Alberta*. Unpublished doctoral dissertation, University of Alberta.
Rice, A.W. (1978). *Individual and work variables associated with principal job satisfaction*. Unpublished doctoral dissertation, University of Alberta.
Richards, D.M. (1971). *Availability and requirements for teachers in Alberta: 1971-1981*. Unpublished doctoral dissertation, University of Alberta.
Richert, G.E. (1968). *Bureaucratic characteristics in educational organizations and their relationship to the leader behavior of the superintendent*. Unpublished doctoral dissertation, University of Alberta.
Robinson, N. (1966). *A study of the professional role orientations of teachers and principals and their relationship to bureaucratic characteristics of school organizations*. Unpublished doctoral dissertation, University of Alberta.

Rourke, M.E. (1983). *The relationship of school effectiveness to administrative structure, professional roles and work interaction of teachers*. Unpublished doctoral dissertation, University of Alberta.

Sackney, L.E. (1976). *The relationship between organizational structure and behavior in secondary schools*. Unpublished doctoral dissertation, University of Alberta.

Sadighian, M. (1975). *A comparative analysis of university goals and governance in North America and Iran*. Unpublished doctoral dissertation, University of Alberta.

Sarros, J.C. (1986). *Educator burnout and its relationship to job satisfaction, role clarity, and job challenge*. Unpublished doctoral dissertation, University of Alberta.

Scharf, M.P. (1967). *An investigation of the relationship between professional role orientation and the social structure of teacher groups*. Unpublished doctoral dissertation, University of Alberta.

Scott, S.H. (1973). *The development of secondary education in Jamaica since independence 1962-1972*. Unpublished doctoral dissertation, University of Alberta.

Sinclair, G.W. (1978). *The development of a program in moral reasoning for educational administrators*. Unpublished doctoral dissertation, University of Alberta.

Stryde, S.J. (1973). *The development of an instrument for describing dimensions of the teaching-learning process*. Unpublished doctoral dissertation, University of Alberta.

Sullivan, K.C. (1976). *Community schools: An analysis of organizational and environmental characteristics*. Unpublished doctoral dissertation, University of Alberta.

Sullivan, T.S. (1987). *Variables related to high school achievement*. Unpublished doctoral dissertation, University of Alberta.

Thorpe, K.M. (1989). *Interpreting the activities of nursing educational administrators: A leadership perspective*. Unpublished doctoral dissertation, University of Alberta.

Toombs, W.N. (1966). *An analysis of parliamentary debates on federal financial participation in education in Canada, 1867-1960*. Unpublished doctoral dissertation, University of Alberta.

Tymko, J.L. (1979). *Accreditation of Alberta senior high schools: A case study of public policy implementation*. Unpublished doctoral dissertation, University of Alberta.

Valentine, P.E.B. (1988). *A hospital school of nursing: A case study of a predominantly female organization*. Unpublished doctoral dissertation, University of Alberta.

Van der Linde, P. (1989). *Preparation, recruitment, selection, and career patterns of directors of education in Saskatchewan*. Unpublished doctoral dissertation, University of Alberta.

Warren, P.J. (1962). *Financing education in Newfoundland*. Unpublished doctoral dissertation, University of Alberta.

Williams, L.T. (1986). *Developmental patterns of teaching careers*. Unpublished doctoral dissertation, University of Alberta.

Williams, M.-J. (1981). *Organizational stress experienced by teachers*. Unpublished doctoral dissertation, University of Alberta.

Wodlinger, M.G. (1980). *A study of teacher interactive decision making*. Unpublished doctoral dissertation, University of Alberta.

Young, B. (1989). *Not finished yet: The stories of four women's careers in education*. Unpublished doctoral dissertation, University of Alberta.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 4, December 1991, 323-332*

HENRY W. HODYSH
University of Alberta

A Note on History, Educational Policy, and the Uses of the Past[1]

The purpose of this paper is to examine how the historian's values inform the history of educational policy. The problem is explored with reference to Kuhn's understanding of value judgment and theory choice. When applied to a case study of policy-oriented history, these understandings suggest that Kuhn's analysis functions as a heuristic device in assessing the normative dimensions of historical research.

When Santayana (1905) remarked that "those who cannot remember the past are condemned to repeat it" (p. 284), he voiced a concern about the value of history that has engaged the minds of historians since antiquity. Santayana's observation has a special significance for the study of policy, especially when it turns to history to illuminate and inform decision making in education. Although historians have shown exceptional skill in analyzing the dynamics of political and social change in the development of educational policy (Hansot & Tyack, 1982; Silver, 1983, 1990; Warren, 1983), few have stepped back in this process to examine the centrality of values that, as Carr (1987) suggests, prompt not only the questions that historians ask, but the answers they receive (Novick, 1988, pp. 522-572). Such values dominate the historian's analysis of educational policy and as standards or principles of worth set the agenda for the methodological concerns which arise (Fischer, 1980, pp. 19-65; Kaplan, 1964, p. 370). The purpose of this article is to examine how the interpretation of historical events and the values on which it is based inform our knowledge of decision making in education.

The implications of this problem for policy historians can be explored from any number of perspectives, but one of particular interest follows from Kuhn's idea of the paradigm (1970a, 1970b, 1977). When broadly defined, the paradigm may be understood as a conceptual framework or theory that represents the accepted beliefs and traditions held by members of the scientific community (1970a, p. 250). The viability of this concept for the study of history is echoed in Hollinger's (1985) observation that Kuhn's idea of science "speaks to, and for historians as few works of *philosophy* of history ever have" (p. 105; compare with Fischer, 1970, pp. 161-170). Hollinger's remark, when directed to the normative foundations of historical research, suggests new possibilities for clarifying the nature of history as policy analysis.

Henry Hodysh, professor in the Department of Educational Foundations, includes the study of historical theory and the methodology of history among his research interests.

In exploring this hypothesis, I first examine the idea of the paradigm and its significance for historical knowledge, noting in particular the nature of value judgments within the paradigm. Then I investigate the implications of that relationship for policy analysis, using as an illustrative sketch Cubberley's instructive yet controversial approach to the history of education. The argument concludes with a critical assessment of the paradigm as a heuristic device for the study of values in history as policy analysis. How, then, do values influence our interpretation of the past and, in turn, affect our understanding of educational policy?

Paradigms, Values, and History

The significance of values for the study of policy-oriented history might be assessed in relation to Kuhn's understanding of the paradigm. For Kuhn, there are two basic meanings of the term. In one sense it represents the disciplinary matrix or "the entire constellation of beliefs, values, techniques, and so on shared by the members of a given community" (Kuhn, 1970b, p. 175). Included in this understanding are the community's metaphors, symbolic generalizations, and shared commitments. The second and more specific meaning of paradigm is contained within the first and refers to the accepted models or examples that replace the community's explicit rules in the cumulative process of normal enquiry (Kuhn, 1970a, p. 250). Such models, articulated and described within the periodical literature, show members of the community how their research is to be done. They result in a "tacit knowledge" for others in the tradition seeking a solution to similar problems in the field (Kuhn, 1970b, p. 192). Over time the emergence and accumulation of problems that cannot be explained by the usual exemplars may lead to paradigm change and the development of a new tradition (1970b, p. 204).

Within the paradigm values are of particular importance. They are widely shared by members of the community and, in the process of application, influence the choices that they make (Kuhn, 1977, p. 335). For Kuhn, values emerge in the community's problem solving activities, reflecting what might be identified as the group's ideology (Kuhn, 1970a, p. 241). They are, in addition, a largely "unconscious concomitant" of theory choice in the preference of one explanation over another (Kuhn, 1977, p. 336). Although members of the community may share a set of common values, they may differ in their application. "In those situations where values must be applied," Kuhn remarks, "different values, taken alone, would often dictate different choices" (Kuhn, 1970b, p. 185; compare with Carr, 1987, p. 131). The distinction between values and their application within a paradigm is critical to the explanation of events, for "knowing what scientists value, we may hope to understand what problems they will undertake and what choices they will make in particular circumstances of conflict" (Kuhn, 1977, p. 290). Values, then, are integral to the community's tradition and evolve with it (p. 335). As Kuhn observes, they are theory-bound to the paradigm and hence tied to its "cognitive commitments" represented in the terms, phrases, and sentences of the language used by individuals (p. xxi). In this context the paradigm provides the researcher not only with a map, but in relation to the

applied values by which in part it is defined, "some of the directions essential for map-making" (Kuhn 1970b, p. 109; compare with Olafson, 1979, pp. 259-260).

In historical research paradigms as models or exemplars would not be limited in time to a specific cultural-historical period. A range of competing research traditions might coexist with no one enjoying the unequivocal support of all members in the historical community. The paradigm, therefore, does not function as a set pattern of history defined, in Fischer's words, as "a right-or-left-handed spiral, a sine curve, a merry-go-round or a loop-the-loop" (1970, p. 162). It serves primarily as a "flexible empirical device," a methodological postulate helpful in unmasking the structural reformations in historical events.

From this perspective values emerge on two different although not mutually exclusive levels. On one level they are critical to the historian as historian who holds an allegiance to certain values in the light of accepted canons of historical research; while on the other they are central to the historian as analyst who clarifies the conditions surrounding the intention, implementation, and outcome of decision making in education. Their impact on one level, however, is not independent of the other, since values as principles of worth might be viewed as instrumental to further ends (Kaplan, 1964, pp. 393-395). Put in another way, the historian's analysis of educational decision making, that description and explanation of policy-related events, is based not only on the shared ideals or inherent values of historical research, but in the choices which individual historians make. These choices reflect operative or applied values, a normative condition of "values in use" that are part of a means-end, value-oriented continuum (Berkhofer, 1969, pp. 101-107, 124-129).

Of the values shared by the community of educational historians two are of particular importance since they speak directly to history as policy analysis. The first embodies a belief in the value of education and schooling. However it may be expressed, historians of education share this value as part of the community in which they exist. The second accepts the reasoned search for truth in the explanation of historical events. Among others, these values as standards or principles of worth command the assent of virtually all historians of education regardless of their association with either traditional or revisionist schools of historical thought.

When applied in the context of a given paradigm, such values assume a specific meaning with their inherent worth becoming instrumental to the pursuit of particular ends (Kaplan, 1964, p. 393; compare with Hodysh, 1989). The shared value, say, of education may be identified with the concept of schooling or, perhaps, enculturation; whereas reasoned explanation may be associated with dialectic, evolutionary, or possibly functionalist approaches to the interpretation of events. As a cognitive commitment, the translation of values from their inherent to instrumental worth is both functional and contextual (Kaplan 1964, p. 395). This in no way detracts from the expectation for the historian to provide a detailed and balanced account of educational policy in the setting where it occurred, but rather

within that understanding affirms the importance of values as an implicit if not an explicit feature of historical research.

Educational Policy and Historical Knowledge: A Case Study

The influence of value judgments on policy-oriented history might be explored with reference to Cubberley's *Public Education in the United States* (1962), a work whose approach to the values of education and reasoned inquiry has special significance for historians of education (compare with Friedman, 1965; Hodysh, 1984; and Lewis, 1967). The book, which significantly shaped the development of educational historiography as a field of study, may not be a policy analysis in the conventional sense, lacking perhaps the specific focus on multiple causation and a detailed concern with the effects of power and the evolving structure of educational policy indicative of more recent studies (Stearns, 1984, p. 110; Warren, 1983, p. 159); yet it presents an assessment of historical change that makes a policy-related contribution by exploring the impact of social and political forces on educational decision making. The central theme of this pioneering history was "the progressive evolution of the power of the state to standardize public schooling, a tale of reformist battles won and further professional goals to accomplish" (Tyack, James, & Benavot, 1987, p. 13). It might be argued that Cubberley's *Public Education* was among the first comprehensive histories that, in the light of a changing America, addressed policy-type questions about education and government in the context of a wider historiographic tradition.

The Emergence of Tradition

Although Cubberley's approach to educational history has been traced to events in the latter part of the 19th century, its definition was primarily a result of developments that for Cubberley occurred at Stanford University about the turn of the century (Cohen, 1976; Cremin, 1965). His influence, along with that of Paul Monroe at Columbia, "the nation's leading historian of education" (Cohen, 1976, p. 306), contributed to the development of educational history as a field of study in teachers' colleges and departments of education. How this was achieved and the problems that were confronted in the process became the basis for a community of educational historians, many of whom shared a common historical and educational vision.

Both Monroe and Cubberley saw the nature and purpose of education in a particular way. Education in Monroe's *Cyclopedia* (Monroe & Kandel, 1912, p. 294) was regarded as the highest conscious effort of society, the culmination of a social process. This ideal, illustrated with reference to the pedagogical views of prominent orators and philosophers in the western educational tradition, covered a time frame from the classical period onward. A similar view appeared in Cubberley's (1920a) *The History of Education* where he emphasized "a history of the progress and practice of education itself" (p. viii). As Monroe remarked, education was an "institutional attempt to realize in individuals the ideals controlling a given people" (1906, p. vi), and this was to be achieved by the extension of universal schooling.

The cognitive commitments of Monroe and Cubberley were represented in a shared vocabulary evident in the use of concepts such as "social evolu-

tion" and "progress." In *A Text-Book in the History of Education,* Monroe (1905) claimed that the application of an evolutionary direction yields new insights into the work of educational theorists from primitive to modern times. It assists the student in establishing a conception of the meaning, nature, process, and purpose of education that "will lift him above the narrow prejudices, the restricted outlook, the failures and the petty trials of the average schoolroom, and afford him the fundamentals of an evolutionary faith as broad as human nature and as deep as the life of the race" (p. ix). There is an alignment of the interests of the school with "social efficiency" and the belief that historical movements "characteristic of the past ... are working themselves out" in the tendencies of the present (p. 751). Such history, closely related to social and political developments, affirms "the spread of western civilization and the state-school idea over the world" (Cubberley, 1920b, p. viii).

Policy-Oriented History

Cubberley (1962) begins his *Public Education* with prefatory remarks about the educational problems of the day and their connection to the recent past. In expanding the thesis developed in his earlier *Changing Conceptions of Education* (1909, p. viii), Cubberley (1962) examines the impact of political institutions and social forces, presenting, as Stearns (1984) would suggest, a procession of events rather than a sustained historical analysis of organizational context and policy precedents (pp. 99-100). Cubberley's *Public Education*, nevertheless, contributes to an understanding of educational policy in describing and explaining, say, the influence of government and society affecting the development and outcome of curriculum change.

A central feature of policy analysis evident in *Public Education* is the explanation of political processes underlying the attention of government and society to education (Silver, 1983, p. 250). To some extent this concern is addressed where, among other examples, Cubberley examines the "struggle" of Horace Mann against the forces for sectarian control of education (1962, p. 298), the impact of the Cousin Report on centralized state educational control (pp. 257-258) and later the efforts of the federal government to establish a Commissioner of Education (pp. 739-740). Current histories of educational policy, however, more fully explore a wide range of variables and multiple causes for historical phenomena, including ideas of state, legislation, power, and their significance for knowledge, social values, and the conditions of learning and achievement (Silver, 1983, p. 250). Although Cubberley's approach to these phenomena lacks in-depth analysis, it reflects the wider implications of government involvement in education evident in his discussion of illiteracy in adult education and the effects of legislation both on industry and child-labor regulations (1962, pp. 570 ff.).

Cubberley's (1962) history ends with a comment on existing challenges in education, including among his topics the influence of child health and welfare developments, new directions for educational change, equality of educational opportunity, and the implications of administrative and professional reorganization. Interestingly, he cites the failure of the individual states to have followed for any length of time "a well-thought-out educational

policy," noting that "a conception of the State as an active and energetic agent for the improvement of educational conditions and the advancement of public welfare has been entirely lacking" (pp. 732-733).

Implications for Analysis

Whether Cubberley's (1962) *Public Education* is viewed as an early form of policy analysis or an educational history that touches on policy problems, it would seem that the function of government in education is among its central concerns. Cubberley's argument, framed on the assumption that social progress is achieved through state control, is grounded in part on normative understandings of education and historical change. As an inherent value, education is defined by the extension of institutional schooling. This process includes not only a linear sequence of development from elementary education onward, but reference to a wider involvement of universities, normal schools, and library and agricultural extension (p. 590; compare with Ward, 1987, pp. 469-472).

Within this paradigmatic framework, Cubberley (1962) establishes the beginning of early education in colonial schools. These origins serve as the basis for an awakening of "educational consciousness" and the eventual triumph of public education in the nation. Concepts of state and legislation at both local and federal levels focus the relationship between education and government in the emergence of decision making. More specifically, the intention and implementation of educational change is described and explained in relation to significant individuals as well as to legislation. Reference, for example, is made to Robert Breckenridge in his efforts to secure public funding for education in the state legislature of Kentucky in the mid-1800s (pp. 187-188). The outcome of decision making is also noted in Cubberley's analysis of curriculum reform as it affected the appearance of the junior high school in California in the early 1900s (p. 555). These problems of policy are explained and articulated with the support of demographic and financial statistics, superintendent and commission reports, memoirs, and a variety of secondary source materials.

Reasoned inquiry as a shared value of historical research is evident in Cubberley's explanation of educational change. This is achieved by means of the genetic or historic method, which contributes to an understanding of the shifting pattern of state educational policy. The method is used not only to establish the origin and sequence of educational development, but to compare educational conditions with what had come before. In tracing the course of events, Cubberley sets out to illustrate how "educational service in its proper setting as a great national institution" has helped to solve many of society's problems (1962, p. viii), an assumption that directs the reasoned description and explanation of historical events to an instrumental end (Berkhofer, 1969, p. 107).

The genetic method influences historical analysis on two levels. The first includes the identification of stages or periods of development and is apparent in the continuity of large-scale change from colonial times to the present. The second level is represented in the small-scale study of government involvement, say, in legislation affecting the consolidation of school

districts in the late 1800s (Cubberley, 1962, pp. 721-727). In both instances, data assist the description and explanation of state control in schooling. Using this method, Cubberley establishes the roots of educational change and, in terms of government policy and direction, the "battles" and "struggles" to achieve social efficiency.

Many have argued, and perhaps rightly, that Cubberley's assumptions of progress, social efficiency and state control are questionable, that his concepts, explanatory techniques, and data are selective and as a result contribute to an exaggerated presentism that has distorted the past. These objections, however, do not diminish the inherent values of education and rationality embedded in his work, but rather address the way in which those values are applied. This change from ideal to operative values is accomplished in Cubberley's decision to consciously or otherwise adopt assumptions and concepts that are instrumental to particular ends. In appreciating the paradigm in which those values are framed, one perhaps might conclude that in knowing what Cubberley as a historian valued, one might hope to understand, but not necessarily approve, the choices that he made (Kuhn, 1977, p. 290).

This illustrative analysis, although limited to a single case study, suggests the following tentative observations:

1. The historian's perception of the past and, therefore, the analysis of educational change is not conceptually neutral. It is structured by a model or a paradigm containing in part values and linguistic categories (Mommsen, 1978, p. 23).
2. These values, such as education and reasoned inquiry, are concurrent with the paradigm and in terms of their application interdependent with the cognitive commitments by which it is informed.
3. As a result, the historian as policy analyst is unable to escape a value-oriented epistemology that influences the criteria of truth and validity that the historian employs (Fischer, 1980, p. 25).
4. This implies that the historian's analysis of educational policy and the conclusions that are drawn are unavoidably theory-bound.

History and Educational Policy: Problems and Prospects

The normative critique of policy-oriented history suggests that an understanding of the past and its possible use in educational decision making requires further analysis of the idea of the paradigm. This elusive concept, subjected to extensive discussion by historians and philosophers alike, has been found wanting, with one critic noting that Kuhn has used it in 22 different ways (Masterman, 1970). Kuhn's attempt at clarification has led to the specific concept of the paradigm as model, a derivative of the generalized term and one that has proved useful in historical research.

It is evident, nevertheless, that difficulties in the use of the paradigm exist not only in relation to its internal components, but in terms of the idea of community. In the community of educational historians, it might be argued that the internecine disputes over the description and explanation of events represent little more than individual efforts to account for historical change. However, the differing approaches, reflecting a variety of paradigms and

ideas of power, may in combination strengthen rather than weaken the historian's understanding of educational policy. And further, whereas historians may disagree on the applied values and assumptions behind the description of policy intention and outcome, they share inherent ideals of education and reasoned inquiry that in part establish a common scholarly tradition.

An additional problem implies that the coincident relationship between the paradigm and its values permits a reductionist interpretation that could remove any meaning these terms might individually possess (Kuhn, 1977, p. 336). The possible resolution of this issue follows a realization that the paradigm as a shared inheritance embraces not only its ideal and applied values, but concepts and assumptions that delineate the paradigm's distinctive character and function. From this perspective values are defined in relation to their companion elements and, it would seem, not reducible to the paradigm as a whole—an awareness that may alert the historian to the particular contribution of those components to an understanding of continuity or change in the educational policy of the past (compare with Fischer, 1970, p. 161).

A final concern centers on the possibility of applying historical knowledge to the formation of present and, perhaps, future educational policy (Graham, 1980, p. 22). In this context the historian employs a process of analogical inference that includes more than an investigation of the origin and nature of decision making in education. Specifically, if values of historical research influence the interpretation of policy-oriented history, might a study of their effect in different historical paradigms contribute to a normative consensus useful in educational planning? This approach, focusing on the clarification of ideal and operative values, assumes that a comparative analysis of historical models might lead to a set of normative guidelines that could inform an educational course of action (Fischer, 1980, p. 49). The study of contrasting paradigms, however, prompts attention not only to problems of identifying similar or equivalent policy-related events from the past, but to the criteria used in selecting for analysis one value-laden model in preference to another.

These challenges, however appropriate, do not dislodge the importance of the paradigm as a heuristic device. They focus, after all, on the function of values in the existential events of the past and, more critically, in the structure of historical interpretation (Scott, 1979, pp. 12-17). In our search for perspective, they contribute to a dialogue within the historical community (Novick, 1988, pp. 626-628) and, in pursuit of policy-oriented research, a determination of its normative problems, if not their solutions. In this context the paradigm is not to be viewed as a rigid framework imposed on historical data from without, but rather, by means of analysis, as a clarification of what in the way of evidence is already there; yet regrettably, some historians have remained indifferent to such theoretical concern. They might agree, however, that many of Kuhn's insights are not foreign to their craft and are based on a knowledge of what their community has long taken for granted. With this understanding, perhaps it is time to reconsider the nature of value judgments and their place in the historical study of educational policy.

Note

1. This article is a revised version of a paper presented to the Canadian History of Education Association Conference in Ottawa, Canada, October 12-14, 1990.

References

Berkhofer, R. (1969). *A behavioral approach to historical analysis*. New York: Free Press.
Carr, E. (1987). *What is history?* Harmondsworth, UK: Macmillan.
Cohen, S. (1976). The history of the history of American education: The uses of the past. *Harvard Educational Review, 46*, 298-330.
Cremin, L. (1965). *The wonderful world of Ellwood Patterson Cubberley*. New York: Teachers College, Columbia University.
Cubberley, E.P. (1909). *Changing conceptions of education*. Boston, MA: Houghton Mifflin.
Cubberley, E.P. (1920a). *The history of education*. Boston, MA: Houghton Mifflin.
Cubberley, E.P. (1920b). *Readings in the history of education* Boston, MA: Houghton Mifflin.
Cubberley, E.P. (1962). *Public education in the United States*. Boston, MA: Houghton Mifflin. (Originally published in 1919)
Fischer, D. (1970). *Historians' fallacies*. New York: Harper and Row.
Fischer, F. (1980). *Politics, values and public policy: The problem of methodology*. Boulder, CO: Westview Press.
Friedman, N. (1965, October). New orders and old historians: Educationists, and the dynamics of academic imperialism. *American Behavioral Scientist*, Reprint IX, 24-29.
Graham, P. (1980). Historians as policy makers. *Educational Researcher, 9*, 21-22.
Hansot, E., & Tyack, D. (1982). A usable past: Using history in educational policy. In A. Lieberman & M. McLaughlin (Eds.), *Policy making in education* (pp. 1-22). Chicago: University of Chicago Press.
Hodysh, H. (1984). T.S. Kuhn's idea of progress and its implications for theory choice in the historiography of education. *Journal of Educational Thought, 18*, 136-151.
Hodysh, H. (1989). Social ideals and the educational purpose of history in the philosophy of William James. *Educational Theory, 39*, 63-69.
Hollinger, D. (1985). T.S. Kuhn's theory of science and its implications for history. In D. Hollinger (Ed.), *In the American province: Studies in the history and historiography of ideas* (pp. 105-129). Bloomington, IN: University Press.
Kaplan, A. (1964). *The conduct of inquiry*. San Francisco, CA: Chandler.
Kuhn, T. (1970a). Reflections on my critics. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 231-278). Cambridge, UK: Cambridge University Press.
Kuhn, T. (1970b). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.
Kuhn, T. (1977). *The essential tension*. Chicago: University of Chicago Press.
Lewis, H. (1967). Bailyn and Cremin on Cubberley and history of education. *Educational Theory, 17*, 56-58.
Masterman, M. (1970). The nature of a paradigm. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 59-89). Cambridge, UK: Cambridge University Press.
Mommsen, W. (1978). Social conditioning and social relevance of historical judgments. *History and Theory*, Beiheft 17, *XVII* 19-35.
Monroe, P. (1905). *A text-book in the history of education*. New York: Macmillan.
Monroe, P. (1906). *Source book of the history of education*. New York: Macmillan.
Monroe, P., & Kandel, I. (1912). History of education. In P. Monroe (Ed.), *A cyclopedia of education* (vol. III, pp. 293-297). New York: Macmillan.
Novick, P. (1988). *That noble dream: The "objectivity question" and the American historical profession*. New York: Cambridge University Press.
Olafson, F. (1979). *The dialectic of action: A philosophical interpretation of history and the humanities*. Chicago: University of Chicago Press.
Santayana, G. (1905). *The life of reason*, Vol. I. New York: Charles Scribner's.
Scott, J. (1979). Reply to the Hilden critique. *International Labor and Working Class History, 16*, 12-17.
Silver, H. (1983). *Education as history*. London: Methuen.
Silver, H. (1990). *Change and the policy process*. London: Falmer Press.

Stearns, P. (1984). History and public policy. In G. McCall & G. Weber (Eds.), *Social science and public policy: The roles of academic disciplines in policy analysis* (pp. 91-128). Port Washington, NY: Associated Faculty Press.

Tyack, D., James, T., & Benavot, A. (1987). *Law and the shaping of public education, 1785-1954*. Madison, WI: University of Wisconsin Press.

Ward, J. (1987). An inquiry into the normative foundations of American public school finances. *Journal of Education Finance, 12*, 469-472.

Warren, D. (1983). The federal interest: Politics and policy study. In J. Best (Ed.), *Historical inquiry in education: A research agenda* (pp. 158-179). Washington, DC: AERA.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 4, December 1991, 333-347*

GRACE V. MALICKY
University of Alberta

Myths and Assumptions of Literacy Education

There has been considerable debate over the past three decades regarding literacy education in the schools. In this article, selected issues are examined in relation to myths surrounding literacy to identify assumptions underlying these issues. From this examination, it is apparent that basic assumptions underlying literacy education have changed very little. Literacy has been and continues to be viewed as an individual achievement, as an educational phenomenon, and as closely linked to economic development and industry.

Introduction

There has been considerable debate over the past three decades regarding the teaching of reading and writing in the schools. As an active participant in some of this debate, I believed that the changes we were debating could have a significant impact on the quality of literacy education. The shift from reading readiness to emergent literacy seemed to have the potential to improve early literacy education for young children. The shift from a skills focus in reading and writing to a process orientation appeared to have even broader implications. However, after three decades of involvement in this debate both as a participant and spectator, the bottom line has changed very little. There are still significant numbers of children being labeled as failures in reading and writing in our schools.

Recently, my work in adult literacy has helped me to reexamine the nature of the debate regarding literacy and I have begun to suspect that while there appear to have been dramatic changes in literacy research and practice, the assumptions underlying literacy education have changed very little over the past three decades. One of the reasons for this involves the mythology surrounding literacy—myths that lead to panic and a corresponding search for panaceas to the "literacy problem." In the first part of this paper, I attempt to expose two of these myths to provide a perspective for the second part, which focuses on the assumptions underlying selected issues in literacy education since the mid-'60s.

Two Literacy Myths

The two literacy myths discussed below are closely related. The first involves the contention that the illiteracy problem is much worse than many people think and is getting worse. The second involves the assumption that literacy

Grace Malicky is a professor in the Department of Elementary Education and Director of the Reading and Language Centre. Her research interests are in literacy development at all age levels.

development is both the cause and the solution to many social problems and will lead to economic development.

The Literacy Crisis

The words *crisis* or *epidemic* frequently appear in both the professional and popular literature on adult literacy. For example Kozol's (1985) book *Illiterate America* is highly alarmist, referring to illiteracy as a "growing crisis" and a "pedagogic time bomb." In Canada, Senator Joyce Fairbairn (1987) spoke in the senate about illiteracy as this country's "hidden shame" and "profound problem."

Two types of statistics are commonly used to support the crisis mentality. The first are statistics on the extent of illiteracy; the second involves estimates of the costs of illiteracy. In regard to the extent of the problem world-wide, Ryan (1985) noted that in 1985 there were 857 million people in the world who were illiterate and that a gradual growth in absolute numbers of illiterates is projected until at least 2000. While the absolute number is increasing, however, there has been and will continue to be a steady decrease in the percentage of the adult population that is illiterate.

The most commonly used statistics in Canada come from census data. The 1986 census data revealed that 3,393,725 persons or 17.3% of the Canadian population aged 15 and over had achieved less than grade 9 in school. In Alberta, the corresponding figures are 192,875 or 10.8% (Statistics Canada, 1989). Census data have frequently been criticized as an indication of literacy (French, 1987; Hunter & Harman, 1979) because it is obvious the "number of school years completed at some time in the past does not define the level of proficiency obtained at that time, nor does it in any way guarantee retention in adult life" (Hunter & Harman, 1979, p. 16). In 1987, Southam News attempted to determine the extent of illiteracy in Canada from a functional perspective. Borrowing items from the National Assessment of Education Progress used in the United States, Southam News drew a random sample of 2,398 persons from across all the provinces and administered both an interview to obtain background information and the functional literacy test. Overall, Southam estimated from their results that 4,502,000 adult Canadians (24%) are illiterate, but the Southam data are at least as problematic as those from the census. The relevance of some test items on the Southam survey have been questioned (Fagan, 1988; Olson, 1988) as have the cut-off point selected to discriminate literate from illiterate adults (Fagan, 1988) and the operational definition of literacy used in the study (Froese, 1988). Using a similar set of items to that in the Southam study, Statistics Canada conducted its own survey with a larger sample in 1990. Results revealed that 16% of Canadian adults do not have sufficient literacy skills to handle everyday reading and writing tasks and that a further 22% are unable to cope with more complex tasks (Brokop, 1991). While these statistics have led to the hysteria noted at the beginning of this section, it is important to keep in mind that both census results and the results of the Southam and Statistics Canada surveys provide only a rough indication at best of the extent of illiteracy in Canada.

The crisis mentality often includes the notion that the illiteracy problem is getting worse rather than better. Graff (1987) refers to this as the myth of decline and argues that the so-called decline has yet to be established over a meaningful time frame. He notes that comparable test results exist only for the past decade and comparisons fail to take into account changes in socio-economic and demographic profiles of student populations in high school and postsecondary settings. Stedman and Kaestle (1987) agree and add that the decline pales when compared to the increase in educational attainment of the population in general over the past 40 years. Finally, Ryan (1985) suggests that "the problem is not a deterioration of reading skills so much as an increase in the levels of education required to participate in an increasingly complex economic and social environment" (p. 380).

The second type of data used to support the crisis argument involves cost estimates. These are based on information that persons with limited education are often the same persons who suffer from one or more of the other major social ills—poverty, unemployment, racial or ethnic discrimination, social isolation, and so forth. (Hunter & Harman, 1979). A causal connection is frequently assumed that Hunter and Harman maintain is reflected in two related literacy myths—first, that literacy is the primary cause of progress, and second, that illiteracy is the cause of poverty and injustice. In the United States, estimates of the costs of illiteracy have been as high as $225 billion yearly in poor job performance, welfare, crime, unrealized tax revenue, and related social ills (Goddard, 1987). Estimates in Canada have been at about $2 billion yearly ("Study Hopes," 1987). While these may be fairly reliable measures of the costs of the social ills outlined, to state that this is the cost of illiteracy per se is clearly an overstatement.

The Power of Literacy

Also overstated are the benefits accrued to literacy. Bhola (1981), for example, indicates that literacy radically changes the way the new literate thinks (*cognitive* outcomes) including such attributes as independence, objectivity, holding individual opinion, capacity for logical analysis, sense of history and universalism. Some of the *social* outcomes of literacy, according to Bhola, include higher income, safety on the job, capacity for increased participation in political institutions, status, and acquisition of modern characteristics such as "keeping informed about the world and taking an active role as a citizen; ... manifesting a sense of personal authority in family, tribe, and sect, and the development of newer nonparochial loyalties; and the concomitant granting of more autonomy and rights to those of lesser status and power, such as minority groups and women" (Bhola, 1981, p. 12). Literacy is also viewed as having *cultural* effects in that the literate individual has improved capacity as a "maker of culture and a creator of knowledge" (p. 12). Finally, Bhola claims that literacy "can make us more fully human" (p. 12).

In contrast to the grandiose statements of Bhola and others, Graff (1987, p. 61) argues that we overvalue literacy and attempts to revise what he refers to as pervasive, uncritical, and outmoded assumptions about the place of literacy in social and individual development. He along with other historians, sociologists and anthropologists (Hunter & Harman, 1979; Levine, 1986;

Olson, 1988) point out that increased literacy is as likely to be an outcome of economic development as a cause of it. From a sociological perspective, Hunter and Harman (1979) state:

> Research suggests that poverty and the power structures of society are more responsible for low levels of literacy than the reverse. For most persons who lack literacy skills, illiteracy is simply one factor interacting with many others—class, race and sex discrimination, welfare dependency, unemployment, poor housing and a general sense of powerlessness. The acquisition of reading and writing skills would eliminate conventional illiteracy among many but would have no appreciable effect on the other factors that perpetuate the poverty of their lives. (pp. 9-10)

Levine (1986) based on his work in Great Britain agrees, stating that the attainment of functional literacy rarely produces outcomes such as employment, personal and economic growth, job advancement, or social integration. In an Edmonton study, Faulk (1987) did not find that the employment status of adults improved as a result of participation in literacy programs.

From a historical perspective, Graff (1987) notes that across time the stratification of literacy has paralleled the stratification of occupations and wealth, and that "the reduction in illiteracy in nineteenth century England ... would appear more as a cultural change wrought by economic growth than as a cause of growth" (p. 66). Anthropological studies have shown that "while literacy was a critical resource, in itself, it did not account for the changes associated with literacy in Western societies" (Olson, 1987, p. 1).

Before leaving this myth regarding the power of literacy, it is important to point out that all uses of literacy may not be desirable. According to Graff (1987):

> literacy and education are better viewed as necessary and not sufficient prerequisites for popular rule and democratic right. The powers of literacy and schooling for control, the negative aspects of education, must be fully recognized. Literacy can and has been employed for social control and for political repression as well. (p. 70)

It appears, then, that we need to question both the crisis myth and the myth that we can solve social problems with literacy programs. According to Sticht (1988), the crisis mentality has spawned many largely unsuccessful quick-fix programs.

Assumptions Underlying Selected Issues in Literacy Education

Reading, Writing, and Literacy

Historically, the term *literacy* was primarily the domain of adult educators but in the late '70s, it began to appear in the literature on the earliest stages of literacy development as educators shifted from a focus on reading readiness to emergent literacy. From a reading readiness perspective, the major focus was on when children were ready to learn to read (Morphett & Washbourne, 1931) or on the relationship of predictor variables such as visual perception, auditory discrimination, and motor coordination with subsequent reading achievement (e.g., deHirsch, Jansky, & Langford, 1966). The focus was on schools, what happened in schools, and on the individual children

themselves with little regard to their broader environment. Reading readiness was clearly conceived as an educational phenomenon.

When reading readiness was reconceptualized as emergent literacy, it was viewed as a continuous process occurring over a long period of time and research changed markedly. However, emergent literacy was still primarily seen as an educational phenomenon as evidenced in Sulzby's (1991) definition of emergent literacy as "the reading and writing behaviors of young children that precede and develop into conventional literacy" (p. 273). The focus of much research in the '80s was on the context of literacy development and on the nature of the interactions between children and significant adults in their lives both in the home and school contexts (Bissex, 1980; Doake, 1981; Juliebo, 1985; Taylor, 1983). The home was considered to be a critical partner in emergent literacy development but the sociopolitical implications of this partnership were recognized by only a few researchers (Taylor, 1983).

The terms *reading* and *writing* still tend to dominate the literature on older children, and these terms clearly have school connotations. The terms imply that reading and writing are essentially educational phenomena and that they have little to do with social or political aspects of children's lives. Indeed, the terms imply that reading and writing are neutral, a highly questionable assumption. As Giroux (1983) and Kretovics (1985) point out, schooling plays a vital role in the reproduction of social and economic relations. This is reflected nowhere more clearly than in the statistics gathered by the Southam News survey (1987) that show that illiteracy is highest among the poor and minorities.

Overall, increased use of the term *literacy* in the literature on children and on what happens in schools has changed little about how literacy is viewed or taught. Literacy continues to be seen as primarily an educational phenomenon and as such essentially neutral. The focus is still on the individual child with literacy development viewed as individual achievement when it is successful and as an individual failure when it is unsuccessful. Partly because of uncritical acceptance of the power of literacy myth, school progress is generally defined in terms of success in reading and writing. It is important to note, however, that this does not reflect "the way the world is" but rather the views of those who have the power to decide "what counts."

Skills or Processes in Reading and Writing

The skills versus processes controversy has included both reading and writing although disenchantment with skills approaches was evident in the literature on reading before that on writing. By the late '60s and early '70s, skills approaches to reading were being viewed as essentially atheoretical and as failing to achieve the North American dream of universal literacy. A decade later, the skills approach to writing was coming under criticism.

Skills approaches are based on the underlying assumption that reading and writing can be taught to children by breaking them down into component skills to be learned one at a time, often through drill and practice. Proctor (1986) notes in relation to writing

> The ideology contained in this curriculum is essentially reductionist; that is, it assumes that the parts are equal to the whole, and that by breaking the

> process of writing into discrete elements learning to write is somehow simplified and, thus, made easier. This is the assembly-line view of language development, wherein the lessons which have been learned from mass production methods used in industry are conceived of as being readily applicable to learning in classrooms. (p. 28)

The influence of industry has also been evident in reading instruction. In the late '60s and '70s, teaching machines reflected this orientation and criterion-referenced programs nearly eliminated the need for teachers much as present-day factories eliminate the need for workers through automation. Teachers became managers of programs rather than teachers of children. This reliance on industry seems related to the close link assumed to exist between literacy and economic development, one of the myths discussed earlier in this article.

The shift from a skills to a reading processes focus did not originate in the '70s, but can be traced back to at least the work of Huey in 1908. However, reading processes were a primary focus of reading research throughout the '70s and '80s (e.g., Kamil, 1984), and this has continued into the '90s as well. A major argument for a process approach was the need to have reading instruction based on an understanding of how children actually read rather than on a priori lists of reading skills deriving from neither theory nor research. The major emphasis has been on learning rather than on teaching, on the child rather than the teacher, and on the how of reading rather than the what. Product has been eschewed in favor of process, and research has tended to focus on the cognitive operations of children as they read. Much of my own research has reflected this orientation (e.g., Malicky & Norman, 1982; Malicky & Norman, 1989; Norman & Malicky, 1987) and my clinical work has been driven by a model of reading processes (Malicky, 1987).

The theoretical framework for this research has come primarily from cognitive psychology, particularly from the area of information processing (Schallert, 1991). Not only have reading educators in the '70s and '80s relied on cognitive theories, but cognitive psychologists have made a major contribution to this research. In fact, cognitive psychology has virtually dominated reading research since the late '60s. The reliance on psychology has, in general, reinforced the focus on the individual and diverted attention away from sociopolitical aspects of literacy.

While cognitive psychology has had some influence on views of the writing process (Flower & Hayes, 1981), the impetus for the shift from skills to process in writing came instead from writers (Murray, 1985) and language arts educators (Calkins, 1983; Graves, 1983). And partly because of the limited attention to writing in elementary schools prior to the '80s, this shift had more influence on classroom practice than that in reading. Process writing appears to provide possibilities for responsibility, freedom, and control for students. Much has been written about children taking ownership of their writing (Graves, 1983; Harste, Woodward, & Burke, 1984). However, although Graves (1983) stressed the need for children to make decisions about their own writing, the authoring cycle or process writing, in practice, has often been used prescriptively. How much freedom do children really have when we require them to "choose topics on their own, rehearse informa-

tion, write, reread what they have written, and revise" (p. 169) every time they write? Innes (1991) recently described a grade 6 classroom where other than the children choosing their own topics, the expectations and activities in a book writing project were the same for everyone. The children all went through the same stages and their work had to be initialled by a teacher after each stage before they were allowed to proceed. To ensure quality of presentation, one teacher typed children's stories, another used a sewing machine to sew pages and the third photocopied stories to facilitate illustrating. This is not unlike the assembly line described above by Proctor (1986).

The trend to a focus on processes has not been without its critics. One of the harshest critics was Hirsch (1987) who argued that the content of instruction ought to be "traditional literate knowledge, the information, attitudes, and assumptions that literate Americans share" (p. 127). He termed this *cultural literacy* and drew up lists of prescribed readings. While Hirsch's ideas are not new, they helped to accelerate the movement toward an increased focus on literature in elementary school programs in the past decade, a move applauded by most language arts educators. However, although "good" literature is preferable to "Dick and Jane," Brause and Mayher (1991) caution that we must be careful that equality and freedom from this perspective do not become "the freedom to be successful by joining the mainstream, established traditions and practices of the elected, powerful governing elite" (p. 260). Bloome (1991) points out that use of good literature raises issues of "who gets to do the choosing, as well as how, when and under what circumstances and what meanings are assigned to those choices" (p. 52). The definitions of what is good literature are those of more powerful groups, and less powerful groups may not be entitled to tell their stories or have their stories included in the literature selected.

Another critic of the process approach has been Delpit (1988) who indicates that process writing deprives children of color of the opportunity to learn what they need to know to gain access to power. She maintains that "Teachers do students no service to suggest, even implicitly, that 'product' is not important. In this country, students will be judged on their product regardless of the process they utilized to achieve it" (p. 287). In a similar vein, Fagan (1990) indicates that "if individuals wish to challenge the present system, they must be able to meet the power brokers on common literacy grounds; otherwise they will be at a disadvantage" (p. 15).

Fagan and Delpit are both referring to the power aspect of literacy that has been predominant in the writing of Freire (1970). But for Freire, power does not involve fitting in with the existing sociopolitical structure, but rather becoming aware of inequities and existing contradictions in society in order to change rather than adapt to them. For him, the what of literacy instruction is critical and ought to originate from people and their circumstances. However, the what would not be the selection of "good" literature by language arts experts. The range of content included in elementary school literacy programs needs to be broader than ever before if one of the goals is to provide equal opportunities for all children. But equal opportunities are not even enough; there must also be a commitment to equality of results. Unless this is achieved, Brause and Mayher (1991) maintain that

"the current practices of segregation and tracking will continue to make a mockery of equality of opportunity" (p. 270).

Overall, both skills and process approaches to literacy education reflect similar views about what literacy is and what it is for. Both focus on the educational nature of literacy and ignore the sociopolitical dimensions. Both focus on the individual. In fact, for years I viewed the child-centered focus of this shift as one of its greatest strengths. But focusing on the child as an individual rather than as a member of his or her social group may isolate the child even further. Possible outcomes of individualized reading include alienation of the child and making him or her even more powerless than before. More and more adult literacy educators are advocating the movement away from the each-one-teach-one format of volunteer tutoring programs to community-based programs in which people work together both to read the word and the world (Freire & Macedo, 1987; Kozol, 1985). The heavy focus of process approaches on the how of literacy detract attention from more critical questions regarding the why and what (Malicky, 1990).

Standardized Tests and Portfolio Assessment

There has been considerable criticism of standardized tests in recent years. Field (1990) notes that test items are selected for their discriminatory power, "that is chosen deliberately to produce failure among some and success among others" (pp. 108-109). In other words, reading failure is a necessary outcome of norm-referenced reading assessment. Brause and Mayher (1991) indicate further that while there have been dramatic changes in assessment over the past decade, demographic data show that these changes have been overwhelmingly counterproductive for the most needy children.

> The limited and limiting tests have emphasized and exaggerated the differences among our students. This has had the effect of emphasizing what they cannot do, rather than what they can do. Further, it has labelled them as potential or actual failures and greased the skids toward their early departure from school. (p. 270)

In lieu of tests, Goodman (1991) has advocated kidwatching, and more recently,the focus has been on portfolio assessment (Mathews, 1990; Paris, 1991; Valencia, 1990). Valencia (1990) indicates that portfolio assessment has the following advantages: it is authentic; it is a continuous, ongoing process; it is multidimensional; and it is collaborative between teacher and student. "It is a philosophy that honors both the process and the products of learning as well as the active participation of the teacher and the students in their own evaluation and growth" (p. 340). However, does portfolio assessment reflect a real change in how literacy is viewed or treated in schools?

One basic aspect that has not changed is that reading and writing are still viewed as individual activities with individual accountability. In relation to writing assessment, Bloome (1991) maintains that

> the change in assessment maintains individual accountability, individual comparison, and definitions of writing instruction as progressive individual skill accumulation. The change in evaluation maintained the former relationships of students, teachers, and evaluators with regard to authority for what determines what counts as good writing and how it should be done. (p. 50)

Field (1990) uses the metaphors *sitting beside to assist* and *sitting above to judge* to elucidate potential differences between portfolio and standardized testing. However, as long as teachers use portfolios to make judgments about children, this potential will not be achieved, and Field points out an even more serious problem. He alludes to the dark side of informal assessment by citing Broadfoot (1984)

> the replacement of formal selective tests and examinations with progressive orientations based on teacher's continuous assessment.... may in practice be a more insidious basis for social control.... thus pupils who for reasons of family background or personality cannot or will not conform to the desired stereotype risk, under such an orientation process, being cooled out long before their potential warrants. (Field, 1990, p. 141 citing Broadfoot)

By appearing to sit beside rather than sit above children, Field (1990) suggests that we may be subjecting children to a more intensive and penetrating form of power and social control. When children put everything they do out in the open, they "let their guards down completely" (p. 191).

It seems likely that with the focus on individual accountability in literacy and with reading in particular still considered to be critical to school progress, the movement to portfolio assessment will do little to solve the problem of inequality for children who come to school from poor and minority homes. There is already pressure on teachers to develop standard criteria for evaluating material included in portfolios. A set of descriptors could facilitate communication about children's literacy development much as we describe their artistic development. However, if these descriptors are correlated with age levels, we will simply be exchanging the norm tables on standardized tests for norm tables based on our opinions. Farr and Carey (1986) state that in the absence of any agreed on list of critical reading or writing skills, we are likely to list the things *we* know. This may lead to even more biased assessment than in the past, and as with standardized tests, we will again create failure and successes through our assessment and by so doing add fuel to the crisis mentality surrounding literacy.

There is little doubt that we need to describe the reading and writing of children to help us plan appropriate instruction to facilitate their literacy growth. However, by using age-related measures, we invite only some children (those who are successful) into the "magical circle of the literate" (Bhola, 1981) or, as Smith (1988) calls it, the literacy club. Others (those who are not successful) are marginalized by our definitions of literacy success and failure. Amoroso (1984) in a study of literacy in a prison context described literacy as a dream of becoming someone for many of the people in his study. Literacy has as much to do with who we are as with what we are able to do. We must build structures into our schools that serve to include rather than exclude children from literate society.

Learning Disabilities: A Focus on the Individual

All three changes discussed to this point focus on the individual, but the individual focus has been even more evident in the movement to learning disabilities as an explanation for reading problems. Most previous explanations had also focused on the child as the source of the problem, for example,

emotional disturbance, mental retardation, neurological dysfunction; the major exception was cultural deprivation. With the advent of the learning disabilities explanation, there was a movement away from compensatory education to remedial programs specifically designed for each child, that is, the individual education plan. The spotlight was removed from society (homes and schools) and responsibility for the reading problem was placed squarely on the child (Coles, 1978) despite the fact that there are many more children who experience reading difficulties from homes of the poor or minorities than from middle-class homes. In fact, the exclusionary clauses of definitions of learning disabilities effectively excluded many minority children from access to remedial reading programs provided through learning disabilities funding.

The learning disabilities explanation was appealing to educators because it defined the child as the one who needed to change to fit the system rather defining the problem as existing in the schools or society at large. Special programs were provided for children but no changes were suggested to the system that provided unequal opportunities and outcomes for the children in the first place.

Coles (1978) suggests that learning disabilities involve providing a medical or biological explanation for a problem that requires a social solution.

> By positing biological bases for learning problems, the responsibility for failure is taken from the schools, communities, and other institutions and is put squarely on the back, or rather within the head, of the child. Thus, the classification plays its political role, moving the focus away from the general educational process, away from the need to change institutions, away from the need to rectify social conditions affecting the child, and away from the need to appropriate more resources for social use toward the remedy of a purely medical problem. (p. 333)

According to Giroux (1987), the label doesn't merely divert attention from the need for social solutions, but also further disadvantages at-risk learners. He is concerned that by having lower expectations for these children, we disempower them further by failing to give them knowledge they need to deal with their everyday lives.

This is not to argue that biological factors are never involved in reading disabilities. Rather, the argument is that by focusing on learning disabilities as residing within children, we lose sight of the fact that this way of assigning meaning and identity to children is socially and culturally determined and as such can be socially and politically changed. We could, for example, consider it acceptable for children at the same age to read and write with different degrees of proficiency much as we expect and accept this in areas such as visual arts and physical education. With less mythology and hysteria surrounding the latter areas, we are willing to accept individual differences and help each child develop at his or her own rate. No one is marginalized; no one is excluded.

The real question for many, though, is whether or not children do learn to read and write better once they have been identified as learning disabled and placed into special programs. This is the major justification for labeling children. However, results of studies on the efficacy of special programs in

resource rooms and special education classrooms have been equivocal at best (Carlberg & Kavale, 1980; Glass, 1983; Haynes & Jenkins, 1986). Rather than questioning the appropriateness of this way of defining children, these results have been interpreted by many as indicating that we have not yet found the best way to teach children once they have been identified. This search for the best way to teach literacy has diverted attention away from the sociopolitical dimensions of learning disabilities, and it is even more pronounced in the learning to read debate.

Wholes versus Parts in Learning to Read

Perhaps no aspect of literacy has been more strongly debated than the beginning stages of learning to read. Chall's (1967) influential book *Learning to Read: The Great Debate* illustrated the polarization of views in the 1960s between those who believed that children should be taught with phonics versus those who supported whole-word approaches. This polarization was recently resurrected by Adams (1990) in *Beginning to Read: Thinking and Learning about Print* with whole language replacing whole-word approaches. Both reached essentially the same conclusion—children need to learn about the letter-sound relationships in written language. Neither addressed the question of why it is important that children learn to read or why more poor children and those from minority backgrounds fail to learn to read than those from middle-class backgrounds no matter what approach is used.

On the surface, whole language seems to reflect a view of literacy quite different from that in bottom-up programs where reading is broken down into subskills to be taught one at a time. Goodman (1986) characterizes language in a whole language program as being meaningful, part of a real event, having purpose for the learner, and being accessible to the learner. He characterizes the learner as having power to use the language. From these characterizations, it would appear that whole language has potential for reaching the emancipatory goals of literacy adopted at the International Symposium for Literacy at Persepolis in 1975. Based on ideas of Freire (1970), literacy was viewed as creating

> the conditions for the acquisition of a critical consciousness of the contradictions of society in which man lives and of its aims; it also stimulates initiative and his participation in the creation of projects capable of acting upon the world, of transforming it, and of defining the aims of an authentic human development. (Hamadache & Martin, 1986, pp. 128-129)

Rather than focusing on these emancipatory possibilities, however, the debate has tended to focus on whether children are taught spelling or phonics in whole language programs. Most children, particularly those from middle-class families, will learn how to read no matter whether we use a phonics approach or a whole language approach because they have been read to many thousands of hours before they arrive at school and hence can make sense of phonics instruction. Alternatively, they may have already figured out much about the sound-spelling patterns of our language as they experimented with writing at home. However, many children from poor or minority homes arrive at school without this advantage and either cannot make sense of phonics instruction because they do not know how written language works or

cannot figure out the relationships between spelling and sound patterns without a great deal more mediation than they generally receive from their teachers.

As with the other issues considered in this article, the whole language debate tends to focus on the view that literacy is an educational phenomenon and on children as individual learners rather than as members of communities. While there tends to be more focus on cooperative learning in whole language than in most traditional programs, assessment practices still make literacy learning an individual, competitive enterprise. The important questions about empowerment and change have rarely been addressed by critics of whole language. As pointed out elsewhere (Malicky, 1990), this may have occurred because whole language has essentially been basalized in Canada. Shannon (1989) notes that the basalization of whole language in combination with the adoption of commercial series by provincial governments has removed control and power from children and teachers and placed it in the hands of publishers and editors. Hence the potential of whole language to meet the emancipatory goals of literacy has effectively been neutralized.

Summary and Implications

My major reason for engaging in the kind of reexamination presented in this article stems from 17 years of clinical work with children and adults defined by schools, society, and themselves as failures in literacy learning. Despite apparent improvements in teaching reading and writing, the situation for these children and adults has changed very little. I began my reexamination convinced that literacy is important to both schooling and life—written language is a convenient and economical way to communicate ideas across time and space. I still believe this, but in our society literacy is viewed as far more than the ability to deal with written language. The crisis mentality and an unquestioned belief in the power of literacy have led to a focus on the how of literacy education—how to teach reading and writing more effectively—with little attention to questions regarding the goals. Only a few of the myths and issues surrounding literacy have been considered in this article but even from this limited examination, a number of underlying assumptions are clearly evident.

First, the focus in literacy education is clearly on the individual. Stress is placed on individual competition rather than cooperation and on individual gain rather than social justice.

Second, literacy is generally viewed as an educational phenomenon. Because of how overvalued literacy is in society, however, it has as much to do with social identity as with constructing meaning from written language. The belief that literacy is educational in nature is related to a further assumption that literacy is neutral, but there are ample statistics to show that illiteracy is higher among the poor and minorities than the rest of the population. Generally, the sociopolitical dimensions of literacy are not recognized, and when they are there are conflicting views regarding whether the poor and minorities should develop the skills and knowledge of the powerful so they can take their place within the current system (e.g., Delpit, 1988) or whether they should develop skills and an awareness of contradictions in

society so they can bring about sociopolitical change. The argument is often presented that the system won't change so we will have to try to change children and adults to fit the system. This is, in essence, what I have been attempting to do in my 17 years of clinical work, and in doing so I have inadvertently supported a system that has been responsible for creating inequalities in the first place. My work has had little to do with empowering individuals and much to do with legitimizing a system that discriminates against certain segments of the population.

A major culprit has been assessment, but our assessment techniques, are merely a reflection of the myths and assumptions we hold about literacy. Because of the power with which we imbue literacy, we have decreed that all individuals must reach certain levels of literacy at specified age levels. To check that we are achieving this goal, we have selected test items designed to pass some individuals and fail others. In other words, by selecting test items to discriminate among individuals, we have predetermined that a certain percentage of the population will fail. But passing and failing is defined as residing within the individual rather than within the assessment techniques, and this is particularly obvious when we label those children who fail to learn to read and write well as being learning disabled. The tests have done their job; they have removed the responsibility for failure from schools and society and placed the blame for this failure on the individual. Coles (1978) refers to this as a classic case of blaming the victim.

A third assumption reflected in our myths and debates regarding literacy has been a belief in technology and in an industrial model of education. We view literacy as the cause of economic development and industrial models as appropriate for planning instructional programs. Assembly-line models permeate much of what we do, and the learning to read debate had been based on the assumption that there is a best way to teach all children. Better methods are seen as the solution not only for illiteracy, but also for the problems associated with illiteracy. However, as Graff (1987) notes, literacy is not the only problem nor the only solution. Much more comprehensive changes in the sociopolitical structures of society are needed to solve problems of poverty, discrimination, unemployment, welfare dependency and powerlessness.

Freire (1970) suggests that an important first step in social change is conscientization. It is crucial that we become more aware of the myths and assumptions that drive much of what we do in literacy education. Only by becoming aware of the contradictions in what we do will we be able to decide what changes are needed and begin to work toward these changes.

References

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Amoroso, H.C. (1984). *Adult conceptions and assumptions about literacy.* Paper presented at the American Research Association Meeting, New Orleans.

Bhola, H.S. (1981). Why literacy can't wait: Issues for the 1980s. *Convergence, 14*(1), 6-22.

Bissex, G.L. (1980). *Gyns at work: A child learns to write and read.* Cambridge, MA: Harvard University Press.

Bloome, D. (1991). Anthropology and research on teaching the English language arts. In J. Flood, J.M. Jensen, D. Lapp, & J.R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 46-56). Toronto: Collier Macmillan.

Brause, R.S., & Mayher, J.S. (1991). Our students. In J. Flood, J.M. Jensen, D. Lapp, & J.R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 259-272). Toronto: Collier Macmillan.

Brokop, F. (1991). *Single mothers in adult basic education: Their words, their worlds.* Unpublished master's thesis, University of Alberta.

Calkins, L.M. (1983). *Lessons from a child: On the teaching and learning of writing.* Exeter, NH: Heinemann.

Carlberg, C., & Kavale, K. (1980). The efficacy of special versus regular class placement for exceptional children: A meta-analysis. *Journal of Special Education, 14,* 295-309.

Chall, J.S. (1967). *Learning to read: The great debate.* New York: McGraw-Hill.

Coles, G.S. (1978). The learning-disabilities test battery: Empirical and social issues. *Harvard Educational Review, 48,* 313-340.

deHirsch, K., Jansky, J.J., & Langford, W.S. (1966). *Predicting reading failure.* New York: Harper and Row.

Delpit, L. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review, 58,* 280-298.

Doake, D.B. (1981). *Book experience and emergent reading behaviour in preschool children.* Unpublished doctoral dissertation, University of Alberta.

Fagan, W.T. (1988). Literacy in Canada: A critique of the Southam report. *Alberta Journal of Educational Research, 34,* 224-231.

Fagan, W.T. (1990). *Adult and child literacy: Need for a model of reconciliation.* Paper presented at the Canadian Society for Studies in Education, Victoria, BC.

Fairbairn, J. (1987). Illiteracy in Canada. *Debates of the Senate, 131*(30), 597-600.

Farr, R., & Carey, R. (1986). *Reading: What can be measured?* Newark, NJ: International Reading Association.

Faulk, J. (1987). *Attitudes of counsellors and policy makers at Alberta Vocational Centres about the unemployed and unemployment.* Unpublished doctoral dissertation, University of Alberta.

Field, J.C. (1990). *Educators' perspectives on assessment: Tensions, contradictions and dilemmas.* Unpublished doctoral dissertation, University of Victoria.

Flower, L., & Hayes, J.R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365-387.

Freire, P. (1970). *Pedagogy of the oppressed.* New York: Herder and Herder.

Freire, P., & Macedo, D. (1987). *Literacy: Reading the word and the world.* South Hadley, MA: Bergin & Garvey.

French, J. (1987). *Adult literacy: A source book and guide.* London: Garland.

Froese, V. (1988). Literacy in Canada: A research report, 1987. *English Quarterly, 21*(2), 120-126.

Giroux, H.A. (1983). *Theory and resistance in education.* South Hadley, MA: Bergin & Garvey.

Giroux, H.A. (1987). Critical literacy and student experience: Donald Graves' approach to literacy. *Language Arts, 64*(2), 175-181.

Glass, G.V. (1983). Effectiveness of special education. *Policy Studies Review, 2,* 65-78.

Goddard, R.W. (1987). The crisis in workplace literacy. *Personnel Journal, 66*(12), 72-81.

Goodman, K. (1986). *What's whole in whole language?* Richmond Hill, ON: Scholastic-TAB.

Goodman, Y. (1991). Informal methods of evaluation. In J. Flood, J.M. Jensen, D. Lapp, & J.R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 502-909). Toronto: Collier, Macmillan.

Graff, H.J. (1987). *The labyrinths of literacy: Reflections on literacy past and present.* London: Falmer.

Graves, D.H. (1983). *Writing: Teachers and children at work.* Portsmouth, NH: Heinemann.

Hamadache, A., & Martin, D. (1986). *Theory and practice of literacy work.* UNESCO.

Harste, J.C., Woodward, V.A., & Burke, C.L. (1984). *Language stories and literacy lessons.* Portsmouth, NH: Heinemann.

Haynes, M.C., & Jenkins, J.R. (1986). Reading instruction in special education resource rooms. *American Educational Research Journal, 23,* 161-170.

Hirsch, E.D. Jr. (1987). *Cultural literacy: What every American needs to know*. Boston, MA: Houghton Mifflin.

Huey, E.B. (1908). *The psychology and pedagogy of reading*. New York: Macmillan.

Hunter, C. St. John, & Harman, D. (1979). *Adult illiteracy in the United States*. New York: McGraw-Hill.

Innes, M.C. (1991). *Six children's responses to literature*. Unpublished doctoral dissertation, University of Alberta.

Juliebo, M.F. (1985). *The literacy world of five young children*. Unpublished doctoral dissertation, University of Alberta.

Kamil, M.L. (1984). Current traditions of reading research. In P.D. Pearson (Ed.), *Handbook of reading research* (pp. 39-62). New York: Longmans.

Kozol, J. (1985). *Illiterate America*. Garden City, NY: Anchor Press/Doubleday.

Kretovics, J.R. (1985). Critical literacy: Challenging the assumptions of mainstream educational theory. *Journal of Education, 167*(2), 50-62.

Levine, K. (1986). *The social context of literacy*. London: Routledge & Kegan Paul.

Malicky, G.V. (1987). *A reading processes approach to diagnosis and remediation*. Unpublished handbook, University of Alberta.

Malicky, G. (1990). The what and why of school literacy. *Reflections on Canadian Literacy, 8*(4), 136-138.

Malicky, G.V., & Norman, C.A. (1982). Reading strategies of adult illiterates. *Journal of Reading, 25*, 731-735.

Malicky, G., & Norman, C.A. (1989). The reading concepts and strategies of adult non-readers. *Journal of Reading, 33*(3), 198-202.

Mathews, J.K. (1990). From computer management to portfolio assessment. *Reading Teacher, 43*(6), 420-421.

Morphett, M.V., & Washbourne, C. (1931). When should children begin to read? *Elementary School Journal, 31,* 496-503.

Murray, D.M. (1985). *A writer teaches writing*. Boston: Houghton Mifflin.

Norman, C., & Malicky, G. (1987). Stages in the reading development of adults. *Journal of Reading, 30*(4), 302-307.

Paris, S.G. (1991). Portfolio assessment for young readers. *Reading Teacher, 44*(9), 680-682.

Proctor, J.R. (1986). *The effect of teachers' theoretical orientations on the writing produced by grade one students*. Unpublished master's thesis, University of Alberta.

Olson, D.R. (1987). An introduction to understanding literacy. *Interchange, 18*(1/2), 1-8.

Olson, D.R. (1988). Mythologizing literacy. *English Quarterly, 21*(2), 115-119.

Ryan, J. (1985). Some key problems in adult literacy. *Prospects, 15*(3), 375-381.

Schallert, D.L. (1991). The contribution of psychology to teaching the language arts. In J. Flood, J.M. Jensen, D. Lapp, & J.R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 30-39). Toronto: Collier Macmillan.

Shannon, P. (1989). The struggle for control of literacy lessons. *Language Arts, 66*(6), 625-634.

Smith, F. (1988). *Joining the literacy club: Further essays into education*. London: Heinemann.

Southam News. (1987). *Literacy in Canada: A research report*. Toronto: The Creative Research Group Ltd.

Statistics Canada. (1989). *Census Canada 1986: Schooling and major field of study*. Ottawa: Government Printing Centre.

Stedman, L.C., & Kaestle, C.F. (1987). Literacy and reading performance in the United States from 1880 to the present. *Reading Research Quarterly, 22*(1), 8-46.

Sticht, T.G. (1988). Adult literacy education. In E. Rothkopf (Ed.), *Review of research in education* (pp. 59-96). Washington, D.C.: American Educational Research Association.

Study hopes to find illiteracy's cost. (1987, May 12), *Edmonton Journal,* p. D4.

Sulzby, E. (1991). The development of the young child and the emergence of literacy. In J. Flood, J.M. Jensen, D. Lapp, & J.R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 273-285). Toronto: Collier Macmillan.

Taylor, D. (1983). *Family literacy*. Exeter, NH: Heinemann.

Valencia, S. (1990). A portfolio approach to classroom reading assessment: The whys, whats, and hows. *Reading Teacher, 43*(4), 338-340.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 4, December 1991, 349-366*

DAIYO SAWADA
University of Alberta

Deconstructing Reflection[1]

Educational theory as well as educational practice are implicitly dominated by a language frame that structures thought and practice into an input-output *form. In this paper I attempt to identify this frame and the way it has predetermined much of our discourse, historically and in the present day, and offer an alternative frame which if nothing else makes us aware that the* input-output *frame does not necessarily disclose* reality.

> We live today in the age of the electronic computer, and that technology has served to shape much of our thinking about how education can and should proceed. We speak of thinking as "information processing" and teaching and learning as "programming," "assembly," or "debugging." ... Attention is needed, however, as to how the computer as metaphor affects our understanding of the processes of learning and teaching. (Kilpatrick, 1985, p. 1)

Although reflection is by no means a recent cognitive phenomenon, it has become a major focus in recent educational research and practice. Yet because reflection has a long history in various forms of consciousness,[2] the recent surge in interest is likely indicative of current cultural forces, inclusive of technology, that make reflection particularly attractive and perhaps seductively so as a way of conceiving if not resolving educational problems. The introductory quotation, highlighting the computer as a hegemonic metaphor pervading education, gives some indication of the pervasiveness of the computer metaphor. However, what the quote does not show, and indeed may even be a victim of, is the silent yet overpowering constraint that this metaphor brings to our understanding of reflection.

It is the intent of this article to deconstruct the implicitly pervasive presuppositions in the computer metaphor as a way of identifying a profoundly different, if not significant, articulation of reflection and recursion. I wish to present an alternate view not as a replacement for the one embedded in the computer metaphor, but as a way of becoming more aware of meanings that are often difficult to discern in the computer metaphor. With this in mind, consider two other statements about reflection from Kilpatrick (1985):

> We somehow move into another dimension when we reflect on what we have done. In reflecting on our experience, we move out of the plane of our

Daiyo Sawada is a professor in the Department of Elementary Education where he teaches courses in mathematics education and educational research. Recent interests include Continental Philosophy, theory of living systems, literary theory, Japanese aesthetics, game theory, metaphor, narratology, cosmology, and most recently, mathematical pedagogy as a poetics.

> everyday existence. We give meaning to experience by getting outside the system. (p. 6)
>
> Recursion, too, requires that the user step outside the system.... To see how recursion works, one needs to get outside the process itself and, so to speak, look down on it from above. (p. 8)

It seems clear that from these two statements "getting outside the system" is a requirement for either reflection or recursion: If we cannot in some sense get outside our present experience, it would not be possible to reflect upon it. A wide sector of the education community would likely agree that "getting outside the system" does express an important condition for reflection and of recursion. Even the use of rather recent research tools such as journal writing, conferencing, consulting with an outside reader, and narrative inquiry generally can be seen as methods for getting outside our experiences in order to see them in a different way. As I endeavor to show however, the above two statements receive their cogency largely because of the dominance of a particular orientation toward experience, an orientation captured rather succinctly and most recently in the computer metaphor. More generally, however, and more pervasively, this same orientation is identified and analyzed at the level of "human communication" in what Reddy (1979) calls the *conduit metaphor*. Within alternate metaphors, these two statements may appear much less cogent or even false. I refer to the problematic "truth" or "falsity" of these two statements as our *Dilemma*.

In order to explicate our Dilemma may I engage in a little make believe: Suppose with me that these two statements were *theorems* in a mathematical system. As a thought experiment let us assume that these two theorems are false. Assuming coherence in the system, it follows that the falsity of these theorems is necessarily due to falsity in one or more of the *postulates* generating the theory. Let us ask:

1. What might these postulates be?
2. What ontological stance is supportive of these postulates?
3. What alternative ontological stances may be appropriate?
4. What particular postulates could be specified within this alternate ontological stance?
5. What meaning do reflection and recursion have given the alternate postulates?
6. What relations exist between the original postulates and the alternate ones, and what sort of pedagogy is implied by these relations?

Question One: The Suspected Postulate(s)

Kilpatrick (1985) offers a clue:

> One can think of a computer—and a mind—as an information-processing device capable of at least three types of processes: assembly, performance, and control. Roughly speaking, assembly processes translate incoming information into a usable form, performance processes use that information together with information stored in memory to produce some outcome, and control processes manage the sequence. (p. 23)

The above remarks are a succinct and explicit expression of the well-known and pervasively used *Input-Output* (I-O) frame that tacitly as well as explicitly pervades cognitive science and learning theory in particular. According to Iran-Nejad (1990):

> Two tacit assumptions seem to govern academic learning: (a) Learning is internalization of external knowledge, and (b) it occurs under conscious executive control. The first assumption implies that external information is the one and only source of learning, and the second suggests that executive control is the one and only internal source of self-regulation of learning processes. (p. 573)

These assumptions are not an arbitrary concoction of learning theorists or cognitive scientists; rather, they are deeply and tacitly embedded in the very texture of our language and culture. Referring to these assumptions as the conduit metaphor (i.e., the general inclination for us to *frame* and therefore to interpret our experience in terms of an input, a channel or conduit, and an output), Reddy (1979) claims that:

> [the] evidence suggests that English has a preferred framework for conceptualizing communication, and can bias thought process toward this framework, even though nothing more than common sense is necessary to devise a different, more accurate framework. I shall thus be trying to convince you of what may be a disturbing premise: that merely by opening our mouths and speaking English we can be drawn into a very real and serious frame conflict. My own belief is that this frame conflict has considerable impact on our social and cultural problems. If we are largely unable, despite the vast array of communications technologies available to us today, to bring about substantive improvements in human communication, it may well be because this frame conflict has led us to attempt faulty solutions to the problem. (p. 285)

The problem of frame conflict that Reddy identifies is particularly disturbing if we agree with Schön (1979) that frame conflicts are immune to resolution "by recourse to the facts" (p. 256). Even more disturbingly, such stubborn conflicts of perspective are "full of potential for violent contention" and harsh words if not harsh actions.[3] The ongoing conflict between labor and management or labor and government in Canada may largely be a matter of frame conflict, and as long as the frames themselves are never discussed, the conflicts are likely to continue unabated. Canada's constitutional crisis is likely another example of frame conflict. In the case of this article, I am suggesting that our differing ways of understanding reflection are another example.

If this is so, it may be worthwhile to briefly consider Reddy's (1979) demonstration, through an abundance of everyday sentences, of how the conduit metaphor pervades and implicitly frames not only what we say and how we say it, but how we think. Let me provide some examples.

1. "Put your ideas into better words."
 In asking for this we may be thinking that we are only helping to clarify the ideas, but we are also imposing the frame that words are "vehicles" or "containers" for embodying ideas. We then "get our ideas across" more clearly if we choose the most appropriate "delivery vehicles" for doing so.

This frame is the "conduit metaphor" and words themselves are conduits for ideas.

2. "Teachers must implement the curriculum."
 Here the process of implementation is a conduit for delivering the curriculum. The longstanding practice of educational experts "developing" the curriculum which is then "implemented" by teachers is thus framed within the conduit metaphor. Much of the field of "Curriculum Studies" has been framed in this way. Postmodern critiques are largely a way of exposing the nature of this type of framing.
3. "The text conveys the message."
 Here text is conduit to deliver messages. In paradigmatic form, the information is encoded into text, the text carries the information, and the reader or interpreter decodes the text. Much of reading theory is still dominated by this model.
4. "We had a good exchange of ideas."
 Here the conduit metaphor is implicit in the process of exchange. We often talk about a discussion as if it were an "exchange of ideas," and yet how does a mere exchange of information lead to the sometimes ecstatic experiences that are generated in some "lively" conversations? Perhaps more than an exchange of information is happening despite our tendency to frame in this way.
5. Here are a few other examples presented without commentary.
 Schools are the most recent way of delivering education.
 Who *gave* you that notion?
 Is there a *program* for tonight's concert?
 An experimental design is a rigorous means of processing the input so that the output is interpretable with maximum clarity.
 In an educational system, objectives prescribe the input; instruction delivers them; evaluation assesses the output.
 Your poem contains silly thoughts.
 Why did we believe that light required a *medium* in order to propagate?

Reddy (1979, p. 290) summarizes as follows:

1. Language functions like a conduit, transferring thoughts bodily from one person to another;
2. in writing and speaking, people insert their thoughts and feelings in the words;
3. words accomplish the transfer by containing the thoughts or feelings and conveying them to others;
4. in listening or reading, people extract the thoughts and feelings once again from the words.

Reddy (1979) found that "of the entire metalinguistic apparatus of the English language, at least seventy percent is directly, visibly, and graphically based on the conduit metaphor" (p. 298). As such, it is not surprising that the I-O frame seems so "obviously" appropriate in Western thinking. But there is a potential blindness lurking in such obviousness that, as Reddy argued, can produce a kind of pathological dependence in our attempts to understand learning or reflection and recursion, "the conduit metaphor can confuse serious attempts at theory building" (p. 307). Indeed, the metaphor has

become so hegemonic that even 30 years ago McCluhan could so prophetically conclude that "The medium *is* the message."

More sinister is Reddy's (1979) warning that, "the conduit metaphor is leading us down a technological and social blind alley ... with mass neglect of the internal, human systems responsible for nine-tenths of the work in communicating" (p. 310). With this in mind, and realizing that the I-O frame is but a particular instantiation of the conduit metaphor, to say that a system "reflects upon itself" means that the system itself becomes *input* to itself. In order for something to be an input, that something must be outside the system into which it is *put.* Therefore, under the conduit metaphor, it is "obvious" that the system must, as Kilpatrick says, "get outside itself" in order to become input to itself. Reflection under this frame thus necessitates "stepping outside oneself." Because the I-O frame is so pervasive in our society as a metaphor (and in science as a model), the necessity of stepping outside ourselves, like the necessity of Euclid's postulates, seem *self-evident.* Thus the suspected postulate is the I-O frame itself, interpreted as a particularly concrete instantiation of the conduit metaphor.[4]

Question Two: Ontology of the I-O Postulate

Strategically speaking, we were led to the I-O frame as the hidden postulate by imagining that Kilpatrick's two statements were "false theorems" traceable to "false" postulates in a mathematical system. We now seek to trace the origin of the "false" postulate in a deeper "false" ontology.[5]

To say that a mind must step outside itself already acknowledges that there is an inside from which to step outside. The inside must be distinguishable from the outside much as a topological closed curve separates a region into an inside and an outside. The ontology of this stance sees the universe as a collection of distinct "insides" (commonly called *objects,* and human beings could be one such) each of which is separate from its outside (commonly called the *environment*).

This ontology I call the *Separation* ontology. That it is a proper context for the I-O postulate is understandable when it is seen how this postulate provides the infrastructure for *explanation* within the separation ontology: Having begun with the stance that the universe is separated into discrete objects, explaining anything more complex than a single object entails finding some way to *connect* the separated objects. The general form for such connection is the conduit metaphor; thus this metaphor is of necessity pervasive in everyday language. In the realm of science and technology the solution to the connection problem is, as one might guess, the I-O postulate: Any object can be theoretically connected to any other object through some hypothesized input-output process. Such processes are standard fare as well as ubiquitous in all sciences.

It may sound trivial, but the fundamental function of the I-O postulate is to *bridge* the original ontological "cut" made in distinguishing object from environment.

This statement may sound far-fetched and even unbelievable to us. I suggest that such unbelievability is itself an indication of the hegemony of the I-O frame. The perplexing difficulty, however, is that the solution (the I-O

postulate), in being a general frame for *repairing*[6] ontological cuts, is now seen as the the mode for all *repairs,* the way for all *progress,* the way to all knowledge, even the way of speaking about communication. It is now taken for granted that progress is a matter of showing how new connections can be made across the infinity of cuts in the now pervasively separated universe, the so-called *Cartesian split* (see Jones, 1982, especially pp. 51-78).[7] In education, this sort of progress is named *Integration*, which again is an obviously "good" characteristic of any educational process. For example, no one argues against the value of, at some point or some level, integrating different subject areas. Integration is thus seen as a solution to our problems of separation. Is it worth asking how it is that we are so dominated by separation (or specialization) as the only way to move ahead? In this sense, integration may simply be an artifact of our current frame.

As a simple instance of the pervasiveness of the I-O ontology in our everyday thinking consider our bodies. It is "natural" if distasteful for us to think of our basic bodily functioning as intaking food and excreting waste, a simple input-output description. We might even ask, "How else could we describe these actions?" Isn't it obvious that we intake food and output waste? We *ARE* an I-O system. Our stomachs are *inside* and the food is *outside.* In order to survive we must *intake* food and *output* waste. It definitely is not a matter of language; it is a matter of *concrete physical REALITY!* How can I deny this?

I won't deny the truth in the above description. I can, however, offer an alternate interpretation of the *concrete physical reality* mentioned above. Consider the two structures shown in Figure 1. From an Euclidean point of view these two structures are quite different; certainly they are not congruent. Figure A seems to have an *inside.* Indeed the upper opening could be a mouth and the lower orifice could be a rectum. That is, it can be described as an input-output structure. However, without any tearing or ripping or *cutting*, this input-output structure can be transformed into the *topologically equivalent* structure on the right. In looking at this "donut," it seems that what was so obviously the "inside" is no longer inside at all: What was formerly seen to be inside is now simply in the hole of the donut. But to be in the hole is to be outside the donut; the former inside is just as outside as the former outside. So inputting and outputting as a process was just a way of describing a certain kind of presence of things as they moved along a boundary (membrane). And since topologically these forms are equivalent, the former basis for distinguishing inside/outside has completely self-destructed if not deconstructed. The input-output description, then, is just a way of (a frame for) talking about what happens along a boundary. It is not the only frame one could use as biologists Maturana and Varela (1980) reveal. Topologically speaking, then, it is no longer "obvious" that, being equivalent to a donut, we are in reality an input-output structure. It is interesting to note that as a branch of mathematics, topology is a rather recent development. It has to do with things like *neighborhood, proximity, nearness,* words that remind us of *situatedness* rather than *input-output.* But because topological thinking is not hegemonic, we do see ourselves as input-output systems in digestion as well as in cognition and education.

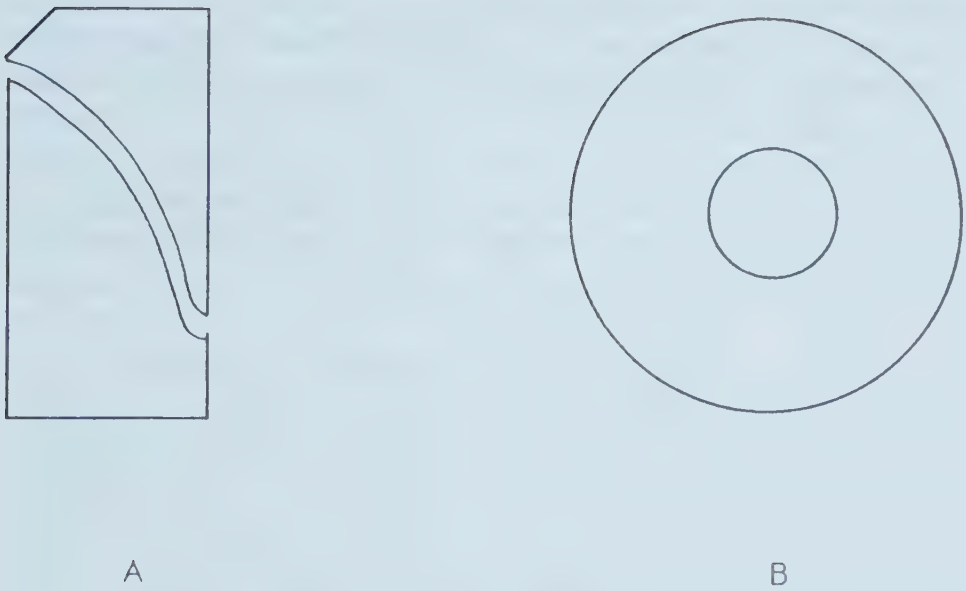


Figure 1. Cross sections of two topologically equivalent 3-D objects.

Question Three: An Alternative Ontology

An alternate ontology is implicit in the I-O postulate (conduit metaphor more pervasively) when it is realized that this postulate is compelling largely because, as suggested earlier, it helps us to re-pair the original *cut* made in the universe. As a device to *heal* the separated universe, the I-O postulate might be more of a technological fix than it is central to a *healthy* universe. In a sense we have become "addicted" to the fix only because we assume that the *natural* state of the universe is to be already cut (Toulmin, 1961). If this natural state is more a product of cultural-lingual circumstances than it is a reflection of any primal state, then a procedure for *bridging* separations would be artifactual at best and debilitating at worst (we would have a world of nothing but artifactual bridges). As an alternative, some attention might be given to a postulate that would not only have us refrain from implementing more cuts as an *automatic* response, but would have us oriented to the ongoing health of an originally given world: Nature would be framed as having an initial integrity not needing *integration*. For lack of a better term, I call this the *Wholeness* Ontology (WO).

WO presumes not only that the universe is already integral, but that wholeness in and of itself is a healthy state (Bertalanffy, 1968). Separating objects from the whole would be seen as an analytical act done by an observer. As observers (farmers, scientists, engineers, teachers, etc.) we have all separated objects. In fact, in this view, identifying objects at all is purely a descriptive act. To the extent that such descriptions destroy the original wholeness, they are necessarily invalid (e.g., diagnosing a child as "a math-phobic grade 4 student with an IQ score of 122, but with short attention span and problems with graphemic but not phonemic reversals, reading 3.2 years below grade level, yet with a self-esteem score in the 89th percentile, etc., etc."). When are such diagnoses complete? How many more cuts are necessary until we can completely "understand" this "disabled learner"? What produced this state of ill health? How many bridges will have to be built so

that this state of disability can bridge itself? Will the child ever be able to crawl out from under the bridges, or does he become these bridges? Is there any child left, or has it become a complicated diagnostic reconstruction of cuts and bridges to span the cuts?

However, descriptions need not destroy the initial integrity. Indeed, descriptions may help us to uncover the phenomenon at a deeper level as in the work of Spencer-Brown (1972). But this will only happen in a conscious way if we are aware that our analyses are products of the frames that produce them. Perhaps a mathematical example is in order.

Although not commonly seen as such, mathematics itself is an example of the making of descriptions that, as creations of the human mind, enhance our understanding but that have no status beyond being creations of our frames. In contrast to this, mathematics in Greek times was seen as fundamentally grounded in *self-evident* truths (Euclid is a prime case). Such self-evidence was provided by the common sense of human experience. Each postulate of Euclid was, as it were, a device for cutting the universe into two parts; one that behaved according to the postulate, and the other that didn't. The part that didn't just didn't exist. Choosing a postulate as a way to cut the universe into two parts is precisely the process of making distinctions as proposed by Spencer-Brown (1972). The postulate as a distinction creates two states, the *marked* state and the *unmarked* state. The marked state is that defined by the postulate; the unmarked state is that excluded by the postulate. For a long time, it was believed that such postulates (the marked states) reflected the *nature of things, the nature of the universe, the self-evident truths of mathematics.* Thus the postulates as descriptions of the nature of things were seen as *reality* captured in logical form. The marked states captured by Euclid's postulates (Euclidean geometry) were thus *reified.* The unmarked states just didn't exist (non-Euclidean geometry). And indeed, for 2000 years, non-Euclidean geometry did not exist, at least not in the minds of people called geometers.[8]

Nowadays in mathematics, the arbitrariness of associating "reality" (common sense) with marked states is revealed as a choice of the mathematician. It is now widely realized that unmarked states can lead to the creation of mathematical systems just as logically valid as those manifested from marked states. Postulates, as choices made by a human, are thus nothing other than the choice of a preferred (self-evident) reality that is then disclosed by that postulate system. The ultimate practicality of mathematics, so eloquently described by Wigner, is a tribute not to the truth of the initial distinctions, but to the coherence of the human mind that chose those distinctions: Having chosen which reality to disclose, should we be amazed when we find that the mathematics thus created actually helps to interpret (disclose) the experiences we have chosen to realize? This is again an example of the power of frames to so powerfully reveal "truth" to the person wielding the frame. WO indicates that we ought not be amazed; surprised perhaps (at the wonderful coherence of being human), but not amazed. Amazement is likely due to believing in the separate existence of mind and reality (the separation ontology); when mind and reality (e.g., mathematics and experience) come together so articulately, we are amazed.

On the other hand, if mind and reality were always together (WO), then we ought not be amazed that they still seem together in our mathematical experience; rather, we may be a little surprised that despite our intricate and elaborate distinction making (our *construction* of mathematics) we have not violated the integrity of the universe (or *multiverse* if we prefer to think in alternative versa)[9] beyond interpretation. From the perspective of WO, if we reify as reality the experiences disclosed by our chosen postulates, our minds may be caught in a Euclidean "mindjacket" too tight to encompass the universe.[10]

Question Four: An Alternative Postulate

While it would be appropriate to enter into the life world of Eastern cultures (Taoism, Zen, Shibusa, Wabi, etc.) to find a postulate resonating with WO, I have chosen to remain within the Western world view, and in doing so I am likely to fall prey to I-O frames implicit in Western languages. Nevertheless, I have chosen to frame my orientation with a set of distinctions that was not available in the East, at least not to the Ancients: The Second Law of Thermodynamics (popularly referred to as the Law of Entropy) that indicates that the "natural" order of the universe is slowly yet inevitably disintegrating (see Rifkin, 1981, for a readable review and Bateson, 1972, for interesting interpretations). If the concept of entropy had been part of the thinking of the Ancients of the East, they may have had second thoughts about existing in harmony with a reality (nature) that was supposedly slowly yet inevitably dissipating. Perhaps they would have been more interested in *enhancing* nature (but of course without violating her). Here in this unhistoric meeting place of East and West, a meeting place that never took place, I see a place for a postulate of enhancement without violation. And yet, taking entropy as a given, is there any hope for enhancement, let alone enhancement without violation? Consider these remarks from Weiner (1967):

> But while the universe as a whole, if indeed there is a whole universe, tends to run down, there are local enclaves whose direction seems opposed to that of the universe at large and in which there is a limited and temporary tendency for organization to increase. (pp. 20-21)

Thus there is hope, at least in "local enclaves," that coherence can be enhanced. Recent work with chemical systems has led to new understanding of the nature of such local enclaves, and I turn now to some of this work as it is significant to an alternate postulate of enhancement. In doing so I caution myself that this work is itself framed within the I-O postulate.

Order Through Fluctuations

The study of the *spontaneous* emergence of order in local enclaves has been motivated in part by the discovery in chemical systems of unexpected patterns in situations of turbulence, particularly in the pioneering research of Ilya Prigogine (1980) who received a Nobel Prize for his work in Dissipative Structures. The name itself suggests that as the given order dissipates (entropy), new structure emerges (enhancement).

Prigogine's study of the spontaneous emergence of new patterns has led him to see universality in the spontaneity of both human beings and the natural world in which they lives. Prigogine and Stengers (1984) state it this way:

> This leads to a new view of matter in which matter is no longer the passive substance described in the mechanistic world view but is associated with spontaneous activity This change is so profound that ... we can really speak about a new dialogue with nature. (p. 9)

This "new *dialogue* of human beings with nature" is a special case of a wider dialogue within nature. This wider dialogue within nature is the setting for local enclaves. In general, local enclaves can be described as special kinds of bounded *open systems*. By this I refer to a system that dialogues, that is, "communicates" *within* its environment, and in this sense is *bounded* yet open to information. For example, in the domain of cognition, such systems could consist of the minds of students communicating within their surroundings (e.g., discussing with other students or with the teacher some aspect of a solution). Though such conversations could be described using the I-O frame, I offer an alternate postulate that may be more appropriate to the conversational setting.

Dissipative Structure theory tells us that of the three types of open systems: (a) those at-equilibrium, (b) those near-equilibrium, and (c) those far-from-equilibrium, only the third type leads *spontaneously* (Prigogine's term) to the emergence of new order. The other types may lead to order, but it will not be an order that emerges spontaneously. When systems propel themselves through reciprocal *recursion* to a far-from-equilibrium state they are subject to spontaneous, dramatic reorganizations of structure. The role of recursion is pervasive in the emergence of this order. Recursion takes the form of *cross-catalytic* and *auto-catalytic* perturbations that, in a compounding (nonlinear) way, drive a system to a highly sensitive state far-from-equilibrium. This is a local enclave, a state in which information not normally sensed can have dramatic (but unpredictable) effects on the new order that emerges. (Unfortunately there is not sufficient space to provide more detail here. For further discussion of Dissipative Structures in education, see Doll, 1986; Sawada & Caley, 1985; Sawada, 1986).

Conversation: An Alternate Postulate

As a paradigmatic embodiment of "dialogue within a local enclave," a *conversation* is a recursive system. The operation of recursion is its essence because it is through recursive interaction that participants in a conversation act as *contextual perturbations* on each other (rather than input into each other) to such a degree that the conversation as a system in its own right becomes sensitive to subtleties of information that recursively lead the conversation to even more subtle configurations far-from-equilibrium. Such a description would, for example, highlight the *immersion setting* for learning a second language rather than the *instructional strategies* themselves.

These kinds of conversational settings are local enclaves to the extent that they can maintain their recursive organization (Maturana & Varela, 1980, 1987; Varela, 1979; Zeleny, 1981). Although it may seem counterintui-

tive, within a conversational setting I take "reflection" as the maintenance of the organization throughout recursive interactions.[11] If enclaves are forced to take on imposed order, then they are forced back to equilibrium, which is the destruction of the enclave.

For example, when an outside "authority" figure (e.g., the principal) enters into the physical space of a lively student conversation, his or her mere presence is often enough of an interruptive perturbation that the conversation disintegrates into a mere *collection of individuals* who now seem just to be "standing around." This would be an example of the destruction, albeit unintentional, of a local enclave; of driving the conversational setting to the equilibrium (entropic) state of a *collection of particles* that just happen to be there (the conversation no longer maintains itself).

The concept of *perturbation* as used above is critical in sensing the difference between the I-O postulate and the Conversation postulate. A conversation is a system of mutually interacting components. Each component (participant in a conversation) "perturbs" other components, not in the sense of "inputting" into other components, but in the sense of *disturbing their equilibrium through a change in the context.* If the perturbations build up recursively to a degree sufficient to drive the components (and possibly the conversation itself) far-from-equilibrium, a new dissipative structure may emerge in the form of new relations among components perhaps at a higher level of equilibrium (a quasi-steady state). The new relations are not in any sense "input"; they emerge as dissipative structures expressing a different, perhaps higher, level of relationship. We might say that the conversation has reached a new plane.

To illustrate these distinctions, consider the many debates that took place in the 1960s and 1970s concerning discovery teaching.[12] While no resolution to the claims and counter claims ever reached a final conclusion, there was one aspect concerning the nature of learning as "personally appropriated knowledge" that was quite controversial. It was claimed that such learning "belonged" to the student. It may be opportune to revive this claim in terms of the conversation postulate.

In discovery learning, information was not *given* (input) to the student; rather, the student was to be placed within a perplexing or ambiguous, or anomalous, or dissonant *setting* (a situation filled with rich perturbations) that required some resolution. In the process of resolving the problem, the student supposedly discovered (perhaps *created* is a better word) some mathematics. It seems to me there are parallels between the discovery learning versus rote learning issue and the conversation postulate versus the I-O postulate. The centrality of the process of perturbation in the conversation postulate is also foundational to the discovery learning dynamic in that the learner is not seen as a being who is just given information to be assimilated (input) so much as a being who is *immersed* in a setting full of perturbations. Generating a higher level of coherence in this situation would correspond to making a "discovery"; it would also correspond to the emergence of order in a *chaotic* situation far-from-equilibrium. In both views, the learner is undergoing a personally appropriated transformation, an example of order emerging within a perturbing context.

A Mathematical Embodiment of Conversation[13]

To this point in our discussion of reflection and recursion the Conversation postulate has been rather amorphous due in part to the limitations of ordinary language. To make these notions somewhat more explicit, as well as to more carefully contrast the Conversation postulate with the I-O postulate, let us consider three mathematical specifications of conversations as systems of recursive functions. Before doing so we need to clarify one item of notation. In indicating the simple sum of say three scalars a, b, and c we normally write a+b+c. In what follows, we use instead the more general notation +/a,b,c in which the three scalars are considered a vector and the addition operation is then defined over the vector. This notation is due to Iverson (1962). Again, I must caution the reader that mathematics too is pervasively affected by the I-O frame.

Consider now a conversation between two participants A and B. Using the above vector notation, and taking the interaction (I) of the participants as the operation, a conversation C_n can be specified jointly as the "married" recursive functions:

$$A_n = I / A_{n-1}, A(A_{n-1}), A(B_{n-1}) \quad (1)$$

$$B_n = I / B_{n-1}, B(B_{n-1}), B(A_{n-1}) \quad (2)$$

In these functions, the conversation is assumed to "chunk" itself into episodes that can be considered as states of the conversation. C_n is the n^{th} state of the conversation while C_{n-1} is the state of the conversation prior to evolving to state C_n. Likewise A_n is the n^{th} state of participant A while A_{n-1} is the state of A prior to evolving to state A_n. Similar remarks hold for B. As well, no initial conditions have been specified because it is impossible to prespecify the original sources of conversations in nature. Some general notions can now be articulated more precisely:

1. A_n is recursively generated within the previous conversational episode C_{n-1} through the perturbational interaction I among A_{n-1}, $A(A_{n-1})$, and $A(B_{n-1})$.

2. $A(A_{n-1})$ is a *composite function* of A with *itself.* As a composite function of itself in a previous episode, A is experiencing itself, A_{n-1}, not so much as an input, but as a *memory.*[14] For example, A could be having "second thoughts" or even arguing with himself or herself. Indeed, mathematically speaking, A_{n-1} is even called and "argument" of A. It is a case of self-argument. As such it may be interpreted as a special case of self-reference, namely *self-awareness.* In other words, in a conversation, self-awareness has a precise specification: It is *self-perturbation* as an argument in a composite function within a recursive interaction. It is a way of experiencing one's experiences by activating memory through self-conversation. In this sense, self-awareness is not so much something that is *input* from the outside as it is the re-cognition of self as recursively present in one's memory.

3. Likewise, $A(B_{n-1})$, as a composite function, is a realization of "other awareness," i.e., B as a perturbation for A. Similarly, $B(A_{n-1})$ is B's awareness of A, in a sense B remembering A from the previous conversational episode. Other awareness is thus *cross-perturbation.* Cross-perturbation is the defining character of conversation as a "marriage" of recursive functions. That is, the "married" aspect of the recursion occurs via the presence

of $A(B_{n-1})$ in equation (1) and $B(A_{n-1})$ in equation (2). Because B_{n-1} in $A(B_{n-1})$ perturbs A and because A_{n-1} in $B(A_{n-1})$ perturbs B, the two functions are explicitly interdependent. It is likely bad enough that in a marriage each partner "perturbs" the other; if each were to directly *input* into each other (or try to) it is not likely that the marriage would last long. More significantly, the "marriage" is a new suprasystem which *is* the *conversation* itself with participants A and B as subsystems (components). The "married" terminology is also used by Hofstadter (1980).

4. Should the cross-perturbational components be trivial (as, for example when A and B are so disinterested they drop out of the conversation or out of the marriage), this can be registered by "dropping" $A(B_{n-1})$ and $B(A_{n-1})$ giving:

$$A_n = I / A_{n-1}, A(A_{n-1}) \quad (3)$$

$$B_n = I / B_{n-1}, B(B_{n-1}) \quad (4)$$

In these "truncated" functions, cross-perturbation is absent and as such the "marriage" is dissolved; the conversation as a suprasystem is no more. We have only two independent recursive functions, that is, participants interested only in themselves (conversational monads). Typically, this would describe a conversation where no one is listening, although of course people may appear as if they were listening. Such conversations rarely lead anywhere. Sometimes teachers talk to children as if listening played no part in the conversation. These truncated equations also describe what happens in a mutually boring conversation: It *self-destructs*. No higher level unity can emerge at the dyadic level; as a conversational dyad, this truncated system is *at-equilibrium*. At the monad level, however, it could reach far-from-equilibrium status.

5. As monads, the equations show that A and B are both self-aware (because of the self-perturbational component). In other words, although they are no longer participants in a conversation with each other (marriage dissolved; no one listening) at the monadic level, they are still in conversation with themselves (perhaps even talking out loud; or, in a problems solving context, thinking aloud to themselves). Thus self-awareness in these equations is precisely "having a conversation with oneself." Of course since the self-perturbational component is also present in equations (1) and (2), self-awareness is an inherent *subproperty* of having conversations with others. Moreover, such self-awareness may be heightened by being immersed in other-awareness.

6. Further truncation leads to:

$$A_n = I / A_{n-1} \quad (5)$$

$$B_n = I / B_{n-1} \quad (6)$$

In these equations there is no perturbation whatsoever, neither self- nor cross-perturbation, since there are no longer composite functions in the model. Instead, we have two individual selves, two independent entities separated from each other. Each is what Ashby (1956) tellingly calls a "regular, determinate, and reproducible" *machine*. They are precisely the kind of entity that is appropriate to what was earlier called the "separation ontology." That is, the I-O postulate is what is left when perturbation is completely truncated from the mathematical specification of conversation:

We are left with an I-O machine *separated* both from itself (no self awareness) and from its environment (no other awareness). For such a machine to reflect upon itself, it indeed must step outside itself and become input to itself. It is an *object* of the kind often created and celebrated in the separation ontology.

7. If one were to program (5) and (6) for computer simulation, the programs would be of the type that *call themselves.* In contrast, if systems (1) and (2) or systems (3) and (4) were programmed, these would be programs that not only call themselves, but also *modify* themselves (self- and other-awareness through self- and cross-perturbation). As such, these *self modifying* programs could lead to unexpected patterns (insights) that are *not* determined by the initial parameter settings. There is a sense of *spontaneity* in such programs which, as Hofstadter (1980) says:

> suggests that suitably complicated recursive systems might be strong enough to break out of any predetermined patterns. [Ashby's "regular, determinate, and reproducible" machines] ... Instead of just considering programs composed of procedures which can recursively *call* themselves, why not get really sophisticated and invent programs which can *modify* themselves—programs which can act on programs, extending them, improving on them, generalizing them, fixing them, and so? This kind of "tangled recursion" probably lies at the heart of of intelligence. (p. 152)[15]

Hofstadter urges us to "get really sophisticated and invent programs which can *modify* themselves." Although his advice is well taken, we should also recognize that, if we take the conversational postulate seriously, then as cognitive systems in conversation we *are* "tangled linguistic recursion."

8. Self- and cross-perturbational processes are central in the mathematical specification in distinguishing the Conversation postulate from the I-O postulate. These two processes are the conversational isomorphs of Prigogine's (1980) "auto-catalysis" and "cross-catalysis", which are so important in catalyzing the *nonlinear* forces that drive a system far-from-equilibrium. Isomorphically, self- and cross-perturbation play the same role in conversational systems.

9. The mathematical embodiment presented here is rudimentary; at best it is only suggestive. For example, the specification as it stands is not *computable* in the sense of Davis (1974). Nonetheless, the models might be helpful in bringing a degree of specificity to the discussion.[16]

Question Five: What Meaning Do Reflection and Recursion Have in the Conversational Postulate?

As has been noted, the self- and cross-perturbations in the recursive interactions of a conversational system generate a far-from-equilibrium enclave in which new order often emerges as dissipative structures. Within this enclave, reflection is simultaneously a process and a product. Let me try to be clearer about this. As a product, reflection can be specified as the new order that emerges within a conversation in just such far-from-equilibrium enclaves. In other words, reflection is generated by the recursive self- and cross-perturbational processes: As the recursion drives the conversation to states that are farther and farther from equilibrium, each state is an enclave in which new

order may emerge in the form of new relations that, by their coherence maintain and stabilize the new quasi-steady state.

Simultaneously, as an ongoing process, reflection is the maintenance of the new order in the form of a *network of relations* that define the new order at a higher level. Thus reflection is simultaneously the new order and the maintenance of the new order.

Even a single cycle of recursion produces a setting for reflection; however, such a setting is likely to be "at- or near-equilibrium" and is therefore not likely to spontaneously give rise to any new order. Most conversations in school, since they are predetermined by a preset agenda (e.g., a teacher-specified objective), are of this kind. However, if the recursion is allowed to drive the conversation far from its starting point, the more likely it is to become an enclave in which new order will emerge as reflection. Unfortunately, in many classrooms, and particularly those dominated by the conduit metaphor, such far-from-equilibrium enclaves, if realized, would likely be seen as "off task" and therefore quickly and decisively brought back to equilibrium in a controlled if not hostile fashion. Formal schooling, in keeping students tightly focused on predetermined objectives (input), likely prevents them from actualizing spontaneous far-from-equilibrium reflections. The coming of the National Council Teachers of Mathematics (NCTM, 1989) *Standards,* with emphasis on communication and problem solving, may provide impetus to allow something like the conversational postulate to be seen as a possible alternative to the I-O frame.[17]

As an illustration of near-equilibrium reflection, Bateson (1972) spoke of the punctuations that participants use in an ongoing conversations. For Bateson, these punctuations were sources of what he called second-level learnings (we would call them reflections) that establish particular personality characteristics. For example, in a dyadic conversation, if Person A punctuates the conversation so that Person B is always responding to Person A's contributions then Person B might well emerge, over a period of time [a period of recursions], with a certain dependency. (Some teachers, particularly those who wish to keep a tight "executive" control on all verbal interaction, punctuate their conversation with students in just this manner.) Such punctuations, experienced daily as linear repetitive recursions, are eventually *reflected* in a new order of personality characteristics (dependency); in turn and over time, this new order will be reflected in a higher level social order (a classroom "climate") in which students rarely show initiative. They have been programmed, so to speak. It could be said that a new (meta) level of "law and order" in the classroom has emerged from the conversations. This kind of meta-order has another name: the hidden agenda or the hidden curriculum. Other styles of punctuation could lead to different personality characteristics (e.g., respect for others) and a very different climate (e.g., a place where risk-taking is supported).[18]

Question Six: A Complementarity

If the previous five questions have been adequately answered, though I fear they have not, then it should be reasonable that reflection and recursion imply that a system must "step outside itself" only under the I-O postulate.

Under the Conversation postulate, reflection and recursion are both *internal* operations, with reflection being generated through recursion. Each level of reflection emerges as the new order in the system at that level. In this way reflection and recursion are necessarily *coexistent.*

What is remarkable is that the exact relationship between the inside experience and the outside experience obtains also between recursion and reflection themselves: Recursion generates the processes that internally drive the system to a far-from-equilibrium enclave in which new order emerges as as network of reflection. Such a network becomes part of the enclave that supports the recursion that generated the network. As a product, such reflection can be described by an observer as "output." The ongoing development of such a system, that is, its history as a self-generating network of recursive processes is as a *complementarity* of reflection and recursion. Varela (1979) has developed a mathematical way of specifying such complementarities in which a higher level system (the ongoing development) is produced from the relations of two interdependent lower level systems wherein one system generates the processes (e.g., recursion) that produce the other (e.g., reflection). In this sense, reflection and recursion form a complementarity in which recursion generates the processes that produce new order (reflection) captured as a network of relations (the complementarity) that supports and maintains the recursion that generates it. Using the symbol * (taken from Varela, 1979) this complementarity can be denoted as follows:

* = network / (reflection / recursion).

Similarly perturbation and transformation, as discussed earlier, form a conversational complementarity:

* = conversation / (transformation / perturbation).

Closing Words

I have given a brief outline of an alternate view of reflection and recursion based on the Conversation postulate and have contrasted this view with that based on the I-O postulate (conduit metaphor). Each postulate brings forth different yet non-independent worlds. Together they form a complementarity that I suggest provides both a more comprehensive and a more authentic disclosure of education.

Notes

1. The work reported here was sponsored in part by a research grant from SARAC of the University of Alberta. I wish also to acknowledge the many conversations with Alton Olson and Tom Kieren whose insights were invaluable..
2. The work of Harry Garfinkle at the University of Alberta is particularly insightful in this regard. He has historically traced major shifts in Western consciousness since 1425 and places reflection as a major form of consciousness emerging in the period 1575-1650, making it a rather dated form of cultural/cognitive awareness.
3. I have had a number of people read this article before it was accepted for publication. While some reactions were supportive and encouraging, others were almost antagonistic. These differences in interpretation are in Schön's terms a result of frame conflict. If frame conflict is immune to the presentation of "facts" then it is most difficult to "improve" the article by marshalling further evidence. The problem is one of the barriers to communication and in this sense this article is likely to be a joy to read or a frustrating experience depending on the degree of frame conflict experienced by the reader.

4. As the ideas in this article began to become clearer to me over a number of years, I began to consider the possibility of putting them on paper (conduit metaphor?). However, the more I tried, the harder it got (a case of frame conflict?). As I write this article I am trying to speak across the chasm without building bridges (conduit?). I may be failing miserably as I get entangled in the very dilemma I am trying to clarify. Could it be otherwise?
5. At the risk of overstatement, the problem of a "false" ontology brings to mind the revolutionary developments in early 19th century geometry in which Euclid's "obvious" postulate (the now celebrated "Fifth") was seriously questioned in terms of its uniqueness in accounting for parallelism. As mathematicians slowly and painfully and somewhat blindly came to question their own commonsense as well as that of their colleagues, alternate contexts (ontologies) from which to construe parallelism gave rise to alternate postulates that eventually were seen as appropriate as Euclid's. In a sense, Euclidean geometry was "self-evident" much as the conduit metaphor is today.
6. The notion of *repair* is itself interesting. Literally in means "to pair again," that is to put back together the two parts of a pair—to repair. As a notion of "fixing" things our current notion of repair is traceable to the original "ontological cut."
7. Darcy and Jayaratna (1985, p. 85) express this point rather succinctly: "The partitioning of any world view into the two components of: 'the system of interest' and 'the attendant environment' necessarily requires the imposition of some arbitrary separation. Within this artificial, intellectual (and therefore invisible) boundary, the search is then directed, usually for solutions to the problems as they are now constrained to appear."
8. Penrose (1989, p. 204) makes a similar point: "The fact that Euclidean geometry seems so accurately to reflect the structure of 'space' of our world has fooled us (or our ancestors!) into thinking that this geometry is a logical necessity, or into thinking that we have an innate *a priori* intuitive grasp that Euclidean geometry *must* apply to the world in which we live [emphasis in original]." Self-evident truths can be very devilish indeed.
9. Maturana and Varela (1980) in their cosmology depict the "universe" as but one manifestation of *multiversa.*
10. The British physicist Bohm (1980) warns: "This sort of ability of man to separate himself from his environment and to divide and apportion things ultimately led to a wide range of negative and descriptive results, because man lost awareness of what he was doing and thus extended the process of division beyond the limits within which it works properly. In essence, the process of division is a way of *thinking about things* [emphasis in original] that is convenient and useful mainly in the domain of practical, technical and functional activities.... Being guided by a fragmentary self-world view, man then acts in such a way as to try to break himself and the worlds up, so that all seems to correspond to his way of thinking. Man thus obtains an apparent proof of the correctness of his fragmentary self-world view though, of course, he overlooks the fact that it is he himself, acting according to his mode of thought, who has brought about the fragmentation that now seems to have an autonomous existence independent of his will and of his desire" (pp. 2-3).
11. Reflections in this sense are the contextual perturbations that together are the immersion setting of a conversation. Conversations taken in this way are their own context. Reflection, as context to conversation as well as being conversation per se, maintains the conversation through perturbational action. Often, of course, the maintenance is of a negative sort, for example, an authoritative demand from outside, resulting in the destruction of the conversation. More on this later.
12. Alton Olson (personal communication).
13. The development to be presented here is based on Sawada and Pothier (1986).
14. For a discussion of memory compatible with the Conversational postulate see Witte, Everett-Turner, and Sawada (1991).
15. It is interesting to note that despite Hofstadter's rather revolutionary suggestions, he still talks of "*improving* on them, *generalizing* them, *fixing* them" indicating that even radical thinkers like Hofstadter are still caught in a language appropriate to the I-O postulate. The hegemony of the conduit metaphor is pervasive indeed.
16. The fascinating book by physicist-mathematician Penrose (1989) provides a penetrating discussion of the limits and possibilities of computability in understanding as well as explaining communication.
17. Indeed, I would suggest that if communicative processes are consciously or unconsciously "conducted" within the conduit metaphor, the *Standards* are likely to be another in a series of small technical changes.

18. Alton Olson (personal communication). Furthermore, in the field of drama, the work of Herman (1991, p. 97) regarding turn-taking in the context of the stage ("dramatic dialogue") reveals an interesting case of communicative punctuation: "In drama, not only the identity of participants, and the roles they adopt, but also who speaks to whom, who is not spoken to, the actual path of interaction and its outcome, the relative quantities of speech apportioned to characters, manner and style of speech, alignments and groupings of characters in the exercise of speech, and power relations established interactively through the manipulation of the speech resource and its alternation are as significant as what is spoken."

References

Ashby, W.R. (1956). *An introduction to cybernetics.* London: Chapman & Hall.

Bateson, G. (1972). *Steps to an ecology of mind.* New York: Ballantine.

Bertalanffy, L.V. (1968). *General systems theory.* New York: Braziller.

Bohm, D. (1980). *Wholeness and the implicate order.* London: Routledge & Kegan Paul.

Darcy, B.G., & Jayaratna, N. (1985). Systems closure and enquiry. *Systems Research, 2*(1), 85-88.

Davis, M. (1974). *Computability.* New York: New York University Press.

Doll, W. (1986). Prigogine: A new sense of order, a new curriculum. *Theory Into Practice, 25,* 10-16.

Herman, V. (1991). Dramatic dialogue and the systematics of turn-taking. *Semiotica, 83,* 97-121.

Hofstadter, D.R. (1980). *Godel, Escher, Bach: The eternal golden braid.* New York: Vintage.

Iran-Nejad, A. (1990). Active and dynamic self-regulation of learning processes. *Review of Educational Research, 60,* 573-602.

Iverson, K.E. (1962). *A programming language.* New York: Wiley.

Jones, R.S. (1982). *Physics as metaphor.* New York: New American Library.

Kilpatrick, J. (1985). Reflection and recursion, *Educational Studies in Mathematics, 16,* 1-26.

Maturana, H., & Varela, F.J. (1980). *Autopoeisis and cognition.* Dordrecht: D. Reidel.

Maturana, H., & Varela, F.J. (1987). *The Tree of Knowledge: The biological roots of human understanding.* Boston: Shambhala.

National Council Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

Penrose, R. (1989). *The Emperor's new mind.* London: Oxford University Press.

Prigogine, I. (1980). *From being to becoming: Time and complexity in physical sciences.* San Francisco: Freeman.

Prigogine, I., & Stengers, I. (1984). *Order out of chaos: Man's new dialogue with nature.* New York: Bantam.

Reddy, M.J. (1979). The conduit metaphor—A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284-324). London: Cambridge University Press.

Rifkin, J. (1981). *Entropy: A new world view.* New York: Bantam.

Sawada, D. (1986). Spontaneous creativity in the classroom. *Journal of Humanistic Education and Development, 25*(1) 2-11.

Sawada, D., & Caley, M.T. (1985). Dissipative structures: New metaphors for becoming in education. *Educational Researcher, 14,* 13-19.

Sawada, D., & Pothier, Y. (1986). *Toward a self-referential theory of partitioning.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Schön, D.A. (1979). Generative metaphor: A perspective on problem-setting in social policy. In A. Ortony (Ed.), *Metaphor and thought* (pp. 254-283). London: Cambridge University Press.

Spencer-Brown, G. (1972). *Laws of form.* New York: Julian.

Toulmin, S. (1961). *Foresight and understanding.* New York: Harper & Row.

Varela, F.J. (1979). *Principles of biological autonomy.* New York: North Holland.

Wiener, N. (1967). *The human use of human beings: Cybernetics and society.* New York: Avon.

Witte, D., Everett-Turner, L., & Sawada, D. (1991). Metaphor as method in recovering teacher's tacit knowledge of play. *Play & Culture, 4,* 24-30.

Zeleny, M. (Ed.). (1981). *Autopoeisis: A theory of living organization.* New York: Elsevier.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 4, December 1991, 367-383*

HEIDI KASS
University of Alberta

The Transformation of Curricular Knowledge

School curriculum appears to be evolving in the direction of increased complexity. The number of stakeholder groups in the curriculum decision making process is leading to competing curricular materials, often with conflicting viewpoints on what should be emphasized. The purpose of the present discussion is to identify a number of possible epistemological perspectives that can be taken on the changing nature of curricular knowledge in science and to consider some of the normative characteristics of each. The varieties of knowledge and value perspectives associated with each inform teachers and students how they are expected to think about the subject, the kinds of knowledge that are to be preferred, and to what kinds of contexts the subject can be expected to apply. An epistemological triad of resultant knowledge, procedural knowledge, and required action conceptualizes the relationship among the components of curricular discourse.

School curriculum appears to be evolving in the direction of increased complexity. Competing curricular materials, each with a specific orientation, for example, process-approach science, environmental science, general science, integrated science, science and society, science and technology, have emerged in many parts of the world, with no particular orientation emerging as a necessary orientation. The purpose of the present discussion is to identify a number of possible epistemological perspectives that can be taken on the changing nature of curricular knowledge in science and to consider some of the normative characteristics of each. The proposition that any perspective that is presented on what is known reflects the values orientation of its developers toward that domain of knowledge is explored in the context of the new curricular realities.

The question of what constitutes curricular knowledge in science is becoming more and more intertwined with the practicalities of designing a curriculum that meets with both professional and public acceptance. In earlier times curriculum design was generally left to subject matter experts and textbook writers. In such a context, questions involving the nature of curricular knowledge were rarely regarded as matters for debate. Relationships between school knowledge and various subject fields tended to be viewed as straightforward and by and large self-evident by both teachers and the public. The large funded science curriculum projects of the 1960s—PSSC Physics, CHEM Study, and BSCS Biology illustrate this point of view.

Heidi Kass is a professor of science education and research design in the Department of Secondary Education. Her research interests include cognitive development and science teaching, the nature of curricular knowledge, and personal knowledge frameworks in learning and teaching.

the public. The large funded science curriculum projects of the 1960s—PSSC Physics, CHEM Study, and BSCS Biology illustrate this point of view.

The situation is now changing rather dramatically. The level of public scrutiny of school curricula is increasing and the involvement of numerous special interest groups as stakeholders in the curriculum decision making process is leading to a variety of often conflicting viewpoints on what should be emphasized (Fensham, 1988). This multiplicity of views and preferences is compounded by the rapid advances in knowledge that characterize numerous subject fields and the emergence of new specializations and approaches to inquiry. The scientific and technological fields are a case in point and will provide the specific examples developed in this discussion.

The proliferation of points of view about what to include in contemporary science curriculum is taking place against a backdrop of concern, particularly in North America, that a serious quality problem seems to exist in science education. The so-called hard sciences of physics and chemistry are often singled out. Enrollment data indicate that many students opt out of science and mathematics courses at the first opportunity (Science Council of Canada, 1984). International comparative studies such as those sponsored by the International Association for the Evaluation of Educational Achievement (Connelly, Crocker, & Kass, 1989) reveal that North American high school students are often clustered in the lower third of average test scores in physics, chemistry, and biology.

Many factors affect both educational outcomes and achievement test scores. To discuss such factors here would be a digression from the major purpose of examining the epistemological nature of curricular knowledge as it is reflected in the sciences. It is also not the purpose to recapitulate and summarize here the large body of scholarly work that has been done in the past 40 years or so in the field of curriculum studies. This work is well documented in the literature. Numerous models have been developed to characterize the processes of curriculum development, implementation, and evaluation and to inform practice in these areas.

Discourse on the nature of the epistemological knowledge forms to be represented in school curricula has not been extensive. School curricula are most often formal curricula with explicitly stated goals, objectives, topics, and skills. In today's education setting, competition for a place in the school curriculum takes place within the context of multiple differing perspectives on what would constitute an appropriate basis for choice. Theoretical treatments of the nature of curricular knowledge, on the other hand, tend to approach the question from one preferred epistemological perspective (Aikenhead & Fleming, 1975; Apple, 1979) or point of view. While such discussions provide additional perspectives to be considered, the problem of how to organize and interrelate the various perspectives remains.

It is suggested that numerous varieties of knowledge and the value perspectives associated with each interact both at the explicit and implicit normative levels in any given curriculum and that these messages are communicated to both the teacher and the student. Often such impressions outlast the recollections of the specific facts and principles. These messages inform teachers and students how they are expected to think about the

subject, about what kinds of knowledge are to be preferred, and to what kinds of contexts the subject can be expected to apply.

The discussion identifies nine knowledge forms that have been found to coexist in science curriculum discourse. The normative nature of each is discussed. The emergence of awareness of these knowledge forms in the course of a secondary school science curriculum development project (Jenkins, 1987) is described. Each illustrates the epistemological values held by a particular interest group. These interest groups are portrayed as emphasizing particular kinds of knowledge. The preferred type of knowledge dominates the prime space of the textbooks and instructional materials that are selected and is also the kind of knowledge that is tested. Each perspective represents a different starting point for approaching the nature of science and can lead to curricula that are very different from each other. Just as it is possible for an individual to hold a variety of perspectives on the nature of knowledge, including knowledge in the scientific and technological realm, curriculum can also reflect a multiplicity of perspectives in a coherent organizational framework. It is argued that students should learn explicitly within the sanctions of schooling that there are different varieties of scientific and technological knowledge and that different value perspectives on their role in science-related social and environmental issues can be held.

Curricular Knowledge as Formalized Knowledge

The transformation of scientific knowledge from its doing phases to its codified and reconstructed forms for purposes of communication to others can be summarized as in Figure 1.

The process of codification, standardization, and transmission of scientific knowledge involves many modifications along the way (Ravetz, 1971). Both explicit and implicit criteria for selection and interpretation operate at each step. What is included, what is omitted, and what is emphasized is a function of the orientation of the person making the decision toward the knowledge. The subtle yet powerful influences of human value orientations on the nature and creation of scientific knowledge itself have been widely discussed since the appearance of Kuhn's essay on "The Structure of Scientific Revolutions" (Kuhn, 1970). At each step in the communication process

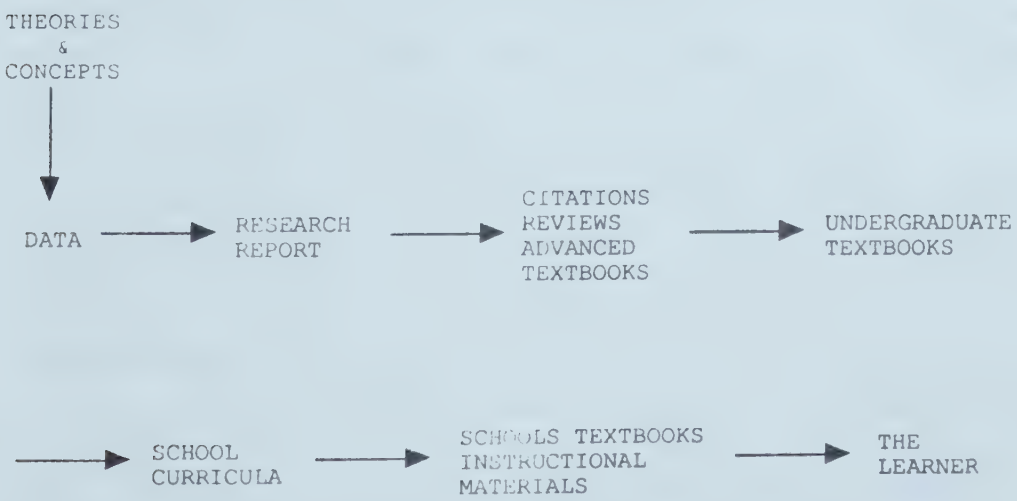


Figure 1. Steps in the standardization of scientific knowledge.

other people more removed from the work engage in decisions affecting its portrayal and place in the whole. Simplification and generalization take place as the material is restated and reorganized, with the inevitable smoothing out of detail and variability. This context stripping and reshaping results eventually in the public view of the scientific enterprise that portrays its development as proceeding in an orderly and linear way to the current organization of standardized facts and principles. The task for the teacher then becomes one of transmitting this body of standardized knowledge that in some sense exists "out there." The main job of the student is to assimilate those particular bits and pieces, problems and examples that are set out in the curriculum of the day (Stake & Easley, 1978; Weiss, 1978). The job of interpretation, of what the concept or fact means or counts for within some conceptual framework or view of knowledge, is either left to the student or discussed in terms of a small number of carefully preselected examples or situations that reflect and reinforce the preferred orientation to knowledge presented in the curriculum.

It needs to be recognized that this process of selection and streamlining has the (perhaps unintended) effect of closing off the subject for some students. The point made by Weisskopf (1976) with respect to physics also applies to chemistry, namely that it is considered inhuman by many students because of its abstract and symbolic concepts that deal with matters under conditions far removed from the human environment and direct human experience. The incompatibility of the picture of science presented in school curricula with the messages received by students from other sources, whether they be the popular media, magazines, or books like Hawking's *A Brief History of Time* (1988) has not escaped the more reflective and well informed of our students.

The nature of the transformations that yield the formalized knowledge of school and university curricula needs to be made explicit to curriculum developers, textbook writers, teachers, and students. Each requires some understanding of this process in order to form a realistic picture of what kind of knowledge one is studying. When these realizations first emerge, some students may be inclined to reject formalized knowledge as being in some sense invalid. From a pedagogic perspective, however, understanding that formal knowledge reflects a high degree of publicly held consensus about what is important in the subject is an insight that may help both the teacher and the student to approach learning in a more productive way.

Normative Perspectives on Knowledge in Science

High school represents the last of formalized science study for approximately four out of five young persons. It is therefore of some importance that school curricula introduce students to the idea of different varieties of knowledge within science and the idea that different value positions about the nature and worth of different kinds of scientific knowledge are held within scientific discourse as well as in science textbooks and classroom discussion.

The pervasiveness of normative elements about the nature of science in science textbooks has been made explicit by Factor and Kooser (1981). Perhaps as a reaction to earlier aggregations of descriptive detail, the science

curricula of the 1960s gave particular value to theoretical knowledge. A number of current curricula, for example, *STSC Chemistry* (Jenkins et al., 1989), *Science Plus* (Atlantic Science Curriculum Project, 1990) place high value on empirical and applicational knowledge. While different textbooks in a field may reflect somewhat different dominant perspectives on the nature of knowledge in the field, it must be remembered that the student is typically exposed to only *one* of these.

How, then, can knowledge in science be viewed in order to present to learners a coherent portrayal of its interactions? The following section presents a set of *nine* knowledge types and some of the different epistemological goals and values that may be ascribed to each. Subsequent sections articulate their influences on what impressions students and teachers may generate about the nature of the subject, the kinds of learnings that are valued, and the kinds of intellectual problems that may emerge when different types of knowledge are inadequately differentiated.

Table 1 presents a set of possible normative perspectives on the nature of knowledge in science. Each perspective, although interacting with one or more others, informs students and teachers about how they are expected to think about the subject, about what kinds of knowledge are to be preferred,

Table 1
Possible Normative Perspectives on the Nature of Knowledge in Science

A. THEORETICAL	- Values theoretical knowledge and theoretical ways of knowing. - Seeks ever greater inclusiveness and predictive power in theory.
B. EMPIRICAL	- Values knowledge gained by empirical means: observations, experimentation, measurement, etc. - Seeks data.
C. APPLICATIONAL	- Values practical and instrumental knowledge and the gaining of knowledge in applied contexts. - Seeks ways of using knowledge, functional knowledge.
D. METHODOLOGICAL	- Values the processes and methods of gaining knowledge. - Seeks problems to solve, establishes criteria for reliability and generalizability.
E. EPISTEMOLOGICAL	- Values knowledge of the limits and validity of knowledge claims. - Seeks to identify underlying assumptions of knowledge claims.
F. SOCIAL RECONSTRUCTIONIST	- Values social, political and economic components of knowledge, social action/reflection. - Seeks emancipatory change, understanding of sociopolitical contexts, and uses of knowledge.
G. PSYCHOLOGICAL	- Values knowledge meaningful to the individual as an active participant in constructing these meanings. - Seeks relevant knowledge, personal knowledge.
H. PEDAGOGICAL	- Values effective and efficient ways of transmitting knowledge. - Seeks to develop literacy, standardization of knowledge.
I. HISTORICAL	- Values developmental characteristics of knowledge over time. - Seeks retrospective understanding of the origin and impact of ideas.

and to what kinds of contexts the substantive and methodological aspects of the subject can be expected to apply. One can ask for each perspective: What kind of knowledge is produced, and how is it valued?

A. Theoretical Knowledge

If current textbooks are any indication, this is the kind of knowledge valued above all else in the sciences. Indeed, Popper (1959) asserts that the empirical sciences are systems of theories. What is meant by theory, and where does theory development begin relative to human experiences and observations (Bradley, 1975) are questions not often addressed in school science courses.

Perhaps in part because theory in chemistry and biology is not nearly as extensively examined in its philosophical foundations as is theory in physics, the complex roles and functions of different kinds of theory in these sciences can become troublesome to students. Part of the puzzlement may be attributed to the tendency for textbooks to ignore such problems. Theory is typically presented as clear-cut and self-evident in its properties as a kind of scientific knowledge. Discipline related and specific problem contexts, if these are considered at all, tend to be viewed as a backdrop for how the theoretical ideas are played out in nature rather than as interacting in a profound way in establishing what the theoretical ideas mean (Kuhn, 1970).

A closer examination of theoretical knowledge in the sciences reveals that the notion of "theory" is itself complex. Examination of science textbooks reveals at least six different ways in which the idea of "theory" is involved.

1. *Grand theory*. This kind of theory, for example, structure of matter, atomic theory, thermodynamics, is concerned with broad explanatory systems. While the context in which the ideas and relationships are used in different branches of science may differ, the desirable attributes of such theory are often implicitly accepted as being the same as in physics, that is, ever greater inclusiveness and predictive power. Individual descriptive contexts are stripped to yield mathematical and symbolic representations wherever possible. Questions having to do with the functions and desirable attributes of such theory within the contexts studied are rarely considered, with the result that students develop little sense of what is involved in formulating a scientific explanation in a specific field.

In the field of chemistry, Hammond and Nyholm (1971) argue that while directed explorations are grounded in theoretical pictures,

> we must recognize that the focus of attention on a single model may seriously inhibit chemical advances. It has been said, and not without a grain of truth, that the factor which distinguishes a chemist from a physicist is that the former argues from 100 weak facts, whereas the latter argues from one strong fact. Without pausing to comment critically on the risks involved in placing too much reliance on one so-called "strong fact," we must concede that chemistry may be a less straightforward subject than physics. The ultimate molecular properties are inferred from indirect measurements, so evidence from many different sources is essential. The emphasis on a single model, may, in fact, distract an investigator from achieving his goal of building better models for molecular behavior. (p.8)

2. *Theory as key constructs* (or concepts), for example, chemical bond, acid. Fensham's discussion of the properties of scientific concepts in terms of intention and extension (Fensham, 1975) is useful here. This is the level at which much of what is conventionally viewed as science is grounded both in its pattern of inquiry and its substantive outcomes. This is also the level at which theoretical ambiguities, for example, what is an acid? and conventions, for example, current flow versus movement of charges, complicate the task of attaching some type of meaning to the theoretical terms.

3. *Theory as prediction/explanation of observation*, for example, the various acid-base theories, types of chemical reaction mechanisms. Much of laboratory work is directed at such a notion of theory. Indeed, the laboratory experiment itself may be viewed as a kind of theoretical entity. Patterns of observation provide meaning both in the sense of extension, that is, the set of objects or events the term denotes and intention, that is, the property or characteristics that is the criterial attribute.

The basic assumption in many science courses is that the students understand what scientific phenomena are. The science curriculum merely provides the explanations. The past 15 years of work on constructivist science learning and teaching (Fensham, 1988; Osborne & Freyberg, 1985; Pines & West, 1985) suggests that the teaching/learning situation is far more complex in its outcomes than has been envisaged. The inability of students to adequately differentiate the theoretical terms from the other kinds of words used (Lemke, 1990) may be attributed to insufficient experience with the content-specific, knowledge-rich nature of laboratory problem solving in science.

4. *Theory as revealed in principles*, for example, Avogadro's hypothesis, "survival of the fittest." Terms such as principle, hypothesis, theory and law are often used loosely and perhaps even interchangeably in scientific discourse. While it may be clear to the textbook writer, curriculum developer, or teacher what these terms mean to them personally, the student may not be able to quickly distinguish this type of theoretical statement from others. The notion of theory as an "unproven idea" is also commonplace in ordinary discourse.

5. *Theories as idealizations*, for example, ideal gas law. Relationships such as the above, which are not universal but idealized reconstructions that come complete with assumptions and limiting conditions, hold a particular place in science. Such theories are not the same kind of logical, or indeed practical, entities as, for example, grand theory.

6. *Theories as models*, for example, electron pair repulsion theory. Correlation between a prior theory and experimental results forms an integral part of science. However, the value of simple and easily used models for appropriate tasks, for example, use of the above model to predict shapes of selected molecules, needs to be placed in context. Such model building helps a great deal in stimulating new experiments. The transient character of many such models is an important aspect of the interaction between theory and experiment in science.

The above relatively lengthy treatment of the varieties of theory in science is intended as illustrative of the broad knowledge category of

"theoretical" knowledge. Consideration of the sense in which each may be viewed as normative within the discipline, within the curriculum, and within instruction requires further analysis.

B. Empirical Knowledge

Here we consider the nature and value placed on knowledge gained by observation, experiment, and measurement. What constitutes scientific data, and on what criteria does one judge its worth? How do the varieties of scientific data pertain to scientific theory?

The science teacher, like the scientist, may experience certain scientific phenomena over and over again in the course of her or his science lessons. The student, on the other hand, may experience the same phenomenon only once, if at all. The complexity of separating out the "scientific observation" from the undifferentiated mass of sense data that the student assimilates is, it is suggested, severely underestimated (Fensham & Kass, 1988). It is also crucial to developing a sense of what is meant by a scientific event (Bhaskar, 1975). As an illustration, it may be difficult for students to understand that even though one can calculate the mass of an electron from its charge to mass ratio and the Millikan oil drop experiment, it is quite another thing to treat it conceptually in the same way as, say, the density of copper.

Several kinds of scientific data may be identified:

1. Results of key experiments that led to advances in theory building and theory validation, for example, Rutherford's alpha scattering. As pointed out by Kuhn (1970) part of the value of such data lies in the historical and/or standardized knowledge realm. Any arbitrarily selected example or instance is not suitable here. It needs to be *the* particular historical one in order to be a part of the shared paradigm of science.
2. Key experiments or observations with implications for theory revision. This includes experiments that have yielded the unexpected, for example, the demise of the idea of the so-called *inert* gases.
3. Normal science—the mass production of scientific knowledge. The results turn out more or less as predicted and any of a range of examples will do.
4. Tricky practical/applied problems, for example, catalysis. Not all common or well-documented scientific observations have a well-understood explanation.
5. Data related to the development and availability of specific instrumentation, for example, spectroscopic techniques, computers. This type of data simply did not exist earlier.

The value placed on data, particularly data experienced by the student is highly problematic in science curricula. In addition to the complex interrelationships with theoretical knowledge, the everyday language of description used in conjunction with scientific data conveys its own set of images. The widespread use of particular adjectives to denote colors is a case in point. Chlorine gas is "greenish yellow," litmus is "red" while phenolphthalein is "pink," and so forth. The student's sense perceptions of the colors may be quite different from the formalized descriptions that are learned.

C. Applicational Knowledge

This type of knowledge places high value on the practical usefulness of knowledge in real contexts and its worth is often contrasted with theoretical knowledge in this regard. The translation of theoretical knowledge to a form suitable for solving particular problems or meeting particular goals may itself be complex. Practical problems typically involve interactions among a much larger number of variables than found in theory-directed problem solving situations. After virtual disappearance in many of the science curricula of the 1960s and 1970s, applied chemistry is slowly being reintroduced in a number of recent school textbooks (Jenkins et al., 1989; Atlantic Science Curriculum Project, 1990).

D. Methodological or Procedural Knowledge

The knowledge inherent in mastery of the techniques and procedures employed in a branch of science is an integral part of that field. Such knowledge is at best only partially described by phrases such as "the scientific method" or "the process of science." As an illustration, the procedural knowledge inherent in knowing *how* and *when* to do a redox titration is *not* the same as knowledge of the generic methods of science. As noted earlier, problem solving in chemistry tends to be more content-specific than in physics and the resulting solutions more localized in their range of application. Casting chemical problems into forms that are potentially solvable and establishing appropriate criteria for reliability and generalizability are a function of ingenuity and practical skill as much as they are application of broad algorithms. When one examines what chemists *do*, the listing of procedures and/or techniques is relatively brief (Fensham, 1983).

The above four varieties of knowledge together form the group that has dominated the sciences and science curricula to date. The five kinds of knowledge to be enumerated in the following sections are more recent in their emergence as curriculum-related perspectives to be taken on knowledge in science. Given the large amount of knowledge in each of the above four types that is available today relative to the amount that can be included in any one curriculum, the value of the kinds of knowledge now outlined may also reside in their usefulness in providing starting points for establishing selection criteria for such inclusion.

E. Epistemological Knowledge Perspectives

The present discussion may be viewed as an exercise in epistemological knowledge. The goal is to identify underlying assumptions of knowledge claims and to develop understanding of the range of applicability, limits, and validity of that which is presented as knowledge in a field. Matters to be considered include what constitutes logic in scientific reasoning, what are the varieties of scientific argument, for example, symbolic, mathematical, and how are these assessed? The need for this type of knowledge to be made explicit in science teaching is stressed by Desautels (1982), who refers to it as a critical view of knowledge.

In the realm of chemistry, students can easily leave their high school courses with impressions such as: all chemical reactions occur very quickly,

all chemical reactions go to completion, all acids are liquids, and other assorted misconceptions. Teachers tend to present only positive exemplars, that is, instances where a particular concept applies. Negative exemplars, also necessary for concept formation (Bruner, Goodnow, & Austin, 1956) are rarely noted. It is, therefore, not surprising that students leave instruction in chemistry with little idea of the range of applicability of the concepts and theories they have learned.

It can be argued that skill in applying epistemological judgments to scientific presentations may be *the* crucial skill in science-related reasoning for the average person to acquire. Social issues having scientific components confront us with increasing regularity. How is one to assess the relative merits of competing scientific positions? Science courses that explicitly deal with such critical reasoning within the context of the other varieties of knowledge can develop such skills in students, incidentally also giving the student a more authentic view of the development of the field (Martin, Kass, & Brouwer, 1990).

Not all theoretical propositions within science have equal "value" in adding to our understanding. Accounts of attempts to improve this understanding and the relative success and failure of various ideas can do much to provide the student with insight into science as a human activity. In one sense all the other perspectives may be viewed as varieties of epistemological perspective.

F. Social Reconstruction Perspective

Understanding of the sociopolitical contexts and uses of scientific knowledge is often mentioned in the context of scientific literacy. As with epistemological perspectives, evaluative dimensions are salient here. Social, political, and economic components of scientific knowledge, for example, areas of inquiry to be developed or closed, scientific research that is classified and/or developed with the support of the military, the ethics of nerve gases, and a host of other issues of the uses and abuses of scientific knowledge, warrant examination. The idea of neutral, value-free knowledge is increasingly seen as a myth. Contemporary science courses need to acquaint learners with some of the current (and historical) dilemmas that have a science component.

G. Psychological Perspective

All of us—including students and scientists—participate in constructing our ideas about the world. This perspective on knowledge values the personal meanings and interpretations held by persons as a result of their participation in the process of construction. Personal relevance and the ability to use the ideas in interpreting specific situations, making decisions, and solving problems are viewed as outcomes of such knowledge. Termed the generative learning model (Osborne & Wittrock, 1985) and based broadly on constructivist theory, the nature of personal science knowledge frameworks and alternative conceptions or explanations are currently active areas of research (White & Tischer, 1986). While the extent of individual capacity to assimilate and compartmentalize inconsistent information has been noted by teachers as well as researchers, its effects on the ability of individuals to integrate such pieces is not well understood. Perhaps what should be studied are persons

who have performed such integration in a functionally productive way, that is, the nature of scientific "talent" or scientific intuition as an adjunct to identifying and describing the various personal conceptions that the population at large may hold.

What attracts individuals to a given field? What are the characteristics of productive collaborations, and how do centers for a particular line of research develop? What factors influence intellectual commitment to a particular theory or problem area? How is the "business" of science, for example, patents, publication, organization of the research team, carried out, and are there differences here in the pure and applied areas?

H. Pedagogic Perspective

As noted earlier, the codification, standardization, and transmission of scientific knowledge involves modifications along the way. What is included, what is omitted, and what is emphasized is a function of the orientation of the person toward the other varieties of knowledge enumerated here. The relative influence of the other perspectives may result in curricula with somewhat different emphasis for the training of research scientists, applied scientists, engineers, other science-related occupations, for example, pharmacy, food sciences, or science teachers. How does the context in which one learns science (pure science, engineering, applied courses, honors thesis, general education, etc.) influence how one presents the key ideas to others—one's choice of sequence, examples, and language used?

Some preferred pedagogic approaches and modifications may also evolve as a result of teaching experience. It is widely accepted that some ideas, for example, atomic theory, are introduced in an historical context while others, for example, reaction mechanisms, are rarely treated in this manner. Some ideas, for example, "diagrams" of atoms, appear in chemistry textbooks primarily as pedagogic rather than as scientific entities. It might be helpful to point out to the learner when such pedagogic simplification and generalization is taking place and the purposes that it serves. The same might be said for the analogies and metaphors that are introduced as aids to description and understanding. Appreciation of the power of a good model or metaphor is an important element in scientific thinking.

Psychological and pedagogic perspectives on knowledge are often interrelated and the psychological can at times act as a barrier to the pedagogic. The portrayal of molecular geometry on the printed page is a common example of such a problem in chemistry teaching, as is the presentation of the generic plant and animal cell in biology. Despite one's best efforts, the conception becomes distorted and its perceived relationship to the actual behavior of these entities vague and inconclusive.

I. Historical Perspective

The understanding of key concepts in science and the language used in conjunction with these concepts may be enhanced by tracing their historic roots. Sometimes the story is "improved" in the retelling, not always to good effect. For example, the history of chemistry is usually portrayed as originating in alchemy. While such an historical emphasis might intrigue the learner, it fails to give account of the extensive applied chemical knowledge that has

existed in human society since ancient times, for example, preserving of foods, fermentation, dyes, medicines, the making of paper, cloth, glass, pottery, and numerous other items. Chemistry has been and to a significant extent remains today a practical science. It is also in this context that problem solving in chemistry is referred to as knowledge-rich task.

Clearly the evolution of science is one of the great adventures of the human imagination. In this sense, some acquaintance with the origins of our key ideas and the men and women who advanced them is a part of our heritage. The revisionary nature of science may also help the graduates of our science courses to understand what it is important to remember and what can be safely forgotten, that is, to develop some sense of intellectual self-knowledge relative to the field.

The Epistemology of Curricular Discourse

The above represents a brief exposition of some of the perspectives that may be taken on knowledge in science. Other possibilities are not ruled out. Each perspective involves value judgments about the kind of knowledge it encompasses and represents a voice in contemporary curriculum discourse. The compilation of the normative perspectives presented here describes a greater range of knowledge production as a part of science than is usually the case. The second set of five normative perspectives underlines the current preference among curriculum developers to view context as a part of the content. It is also recognized that different groups—researchers, graduate students, teachers, undergraduates, high school students, and elementary students—have a qualitatively different experience of the subject. Each begins with a different perspective.

A curriculum presents a learner not only with a perspective of the discipline as it was and currently is, but also of what it will be (or should be). How a particular discipline within science is portrayed as a potential area of science is also viewed as important, given current awareness of the underrepresentation of women and other groups in science-related fields (Sjoberg & Imsen, 1988; Kahle, 1988).

The growth of the scientific enterprise itself is contributing to the making of decisions about what to include, what to leave out, and how it should be organized in an increasingly problematic environment. When such an undertaking is extended to a finite entity such as a school curriculum, the existence of a basis for selection, whether explicit or implicit, that represents a workable degree of consensus among the stakeholders becomes of practical as well as philosophical importance. Programs that are perceived as "too different" from the prevailing orientation receive little support and are quickly replaced. Yet the idea that the focus on the concepts and theoretical principles of a science discipline, while perhaps still the prevailing orientation, is not of itself a necessary orientation, is gaining a foothold in many parts of the world (Fensham, 1988).

How might these perspectives be reflected in contemporary science curriculum in a coherent and pedagogically sensible way? An early interpretation of the question focused on making the dominant orientation explicit. Walker (1971) suggested that the "platform" of curriculum writers be set out

as an explicit part of the curriculum material. Connelly (1972) recommended that any curriculum should present a sufficiently clear exposition of its epistemological basis to permit a clear choice among alternatives. The teacher, school system, or government education ministry can then choose among the available perspectives and implement that particular view in a way appropriate to the characteristics of the setting. The variety of theoretical, applied, environmental, and science-technology-society oriented science programs that are currently in existence or under development attest to the viability of this approach.

Another approach is also possible. It should be pointed out that just as it is possible for an individual to hold a variety of perspectives on the nature of knowledge, including knowledge in the scientific and technological realm, curriculum and instruction can also reflect a multiplicity of perspectives within a coherent organizational framework. No one perspective, no matter how well developed, is adequate in reflecting the plurality of perspectives that can be used in relation to the subject. Indeed, it could be argued that study of one single science subject or specialization is itself not a sufficient vehicle for promoting an awareness of science. We live in a pluralistic world, with numerous alternatives for approaching it and learning to cope with it, and students need to be taught how to examine matters from a number of alternative perspectives. The linear, cumulative, top-down ways of thinking and of representing science are declining in usefulness in the present human context and, at some level, our high school students are aware of this.

A more immediate practical problem is the tunnel vision that quickly sets in when only one perspective is emphasized, whether it be societal, historical, theoretical, or whatever. No one perspective can be clearly designated as the "correct" one. All are right and all are wrong at the same time, but all are useful in that they represent different ways of viewing the subject and different starting points for inquiry. Students can be taught how to approach science with this type of orientation in order to think in an integrated, holistic way about the nature of science. Such perspectives are not mutually exclusive, but interact in many subtle ways.

Can this be done within the framework of a single curriculum? Can this be done to produce a coherent integrated course, or will the product be piecemeal, fragmented, and lacking in academic rigor? Clearly this is a question that cannot be answered in simplistic terms. Various alternative ways of approaching a multiperspective curriculum organization need to be devised, implemented, and assessed. The view that since knowledge is relative hence it does not matter what we teach, is *not* what is being said here. Obviously, it does matter, but what and how to teach those aspects of science that reflect the current nature and importance of the field in ways that are meaningful to more than the college bound in the area requires innovation and exploration.

Roberts (1982) has developed the notion of curriculum emphasis as an approach to developing a multiperspective framework for science curriculum. A program entitled *STS Chemistry* (Jenkins et al., 1989), among others, has used this idea as an organizational rationale in a way that operationalizes the approach for teachers. Various Canadian education min-

istries are also examining this rationale, particularly in the context of science-technology-society oriented curricula (Alberta Education, 1990).

Questions about what perspectives and whose perspectives should prevail, what the value and contextual dimensions would look like at any given point in time, and what contribution the teacher needs to make in response to local demands and changing circumstances will likely be with curriculum developers who are experimenting with various multiperspective approaches for some time.

The Epistemological Triad

A final and perhaps most fundamental question to be considered in the context of the present discussion of evolving curricular complexity involves the contribution that the student makes in developing her or his own understanding of science. How can we educate the learner to respond as sensitively and intelligently as possible to the complexities inherent in doing science, communicating science, and learning science or, for that matter, any other discipline? How does a coherent and functional world view develop?

Both professional experience and countless achievement tests support the hypothesis that a great deal of the science learned in schools is memorized using various rote learning strategies but is not understood by the students. Yet students, like people in general, attempt to make sense of the information they receive from their environment by relating it to knowledge that they already have. The outcomes of these attempts may provide mixed results (Fensham, 1983).

One approach that is meeting with acceptance among science teachers is the development of various metacognitive skills in students in order to increase their self-awareness as learners. The Peel Project (Baird & Mitchell, 1986) in Australia has explored this area in relation to increasing the effectiveness of student learning behaviors with some success.

The question of how students can come to understand—not only to know what it is they know, but also how they come to know it—also emerged within the context of a classroom-oriented approach to the development of a secondary school chemistry program (Jenkins et al., 1981) that questioned many of the traditional views of curriculum and the features of curriculum materials. In the course of a detailed epistemological analysis of the curriculum material as it went through various revisions in the development and classroom trial process (Jenkins, 1987) an epistemological relationship emerged that serves to link the curriculum in practice (i.e., "text" in the general sense) and the students. As currently visualized, the elements in this *epistemological triad* are:

a. *substantive knowledge*—this constitutes the "text" in an epistemological sense;
b. *methodological knowledge*—the method by which the student is to establish the validity of the textual assertion;
c. *action component*—the requirement of the student to respond to the text by producing other knowledge.

The labels for a, b, and c remain problematic. The triad may also be described in terms of (a) pedagogy, (b) process, and (c) epistemology in the

corresponding dimensions above. In relation to science teaching, the three questions considered might be: (a) What is known? (b) How did scientists come to know this? (c) How am I to come to know this, that is, how am I to respond to this by producing new knowledge for myself? Some would label the latter as "coming to understand." Figure 2 portrays some possible linkages in an epistemological structure for understanding a concept in science.

All three elements of the epistemological triad are either explicitly or, more often, implicitly present in curriculum as it is played out in the classroom. This presence is dynamic in nature, characterized by interaction and development. The student's epistemological views both of the epistemological features of what is to be learned and the nature of herself or himself as a learner influence the learning outcome. For example, a student who views science as a collection of neutral facts may interpret the personal action component of a science learning task as one of getting the right answer. A student who views science as being grounded in human experience may interpret the personal action component as one of figuring out how the knowledge is created. Both may be working on the same laboratory exercise or assignment.

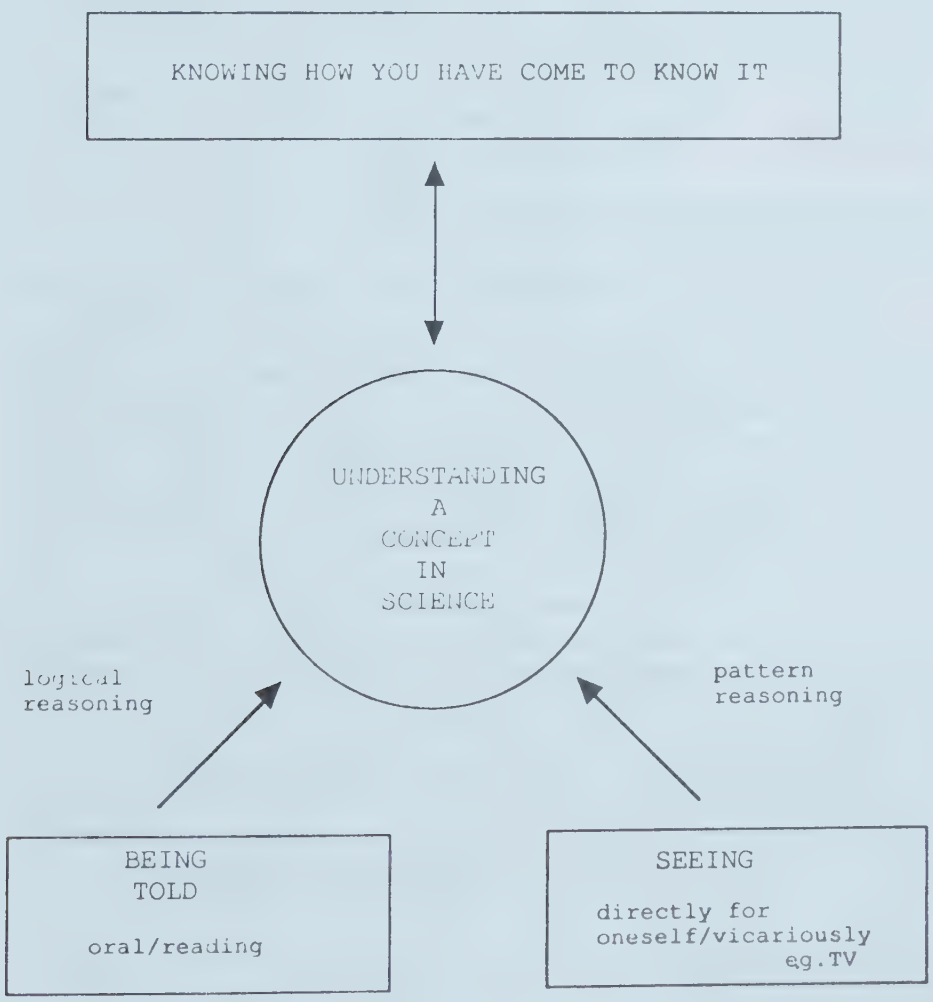


Figure 2. An epistemological structure for understanding a concept in science.

From a curriculum standpoint any curriculum emphasis or normative perspective on knowledge needs to be supported by evidence of epistemological content. As an example, given the emphasis placed on scientific theory and theoretical ways of knowing in science curricula, it is noteworthy that the epistemological question of what ways of knowing are used to derive a theory is rarely addressed in science curricula. Is it a surprise, then, that students fail to see any connection between theoretical science and the "epistemology in practice" of the real world?

Curricular Complexity and Pedagogic Reality

Whether curriculum in science will continue on a path toward increasing complexity or multiplicity of competing emphases and perspectives or evolve to a new metalevel of coherence will ultimately rest, as do most curricular questions, on what this means in the classroom. Some teaching techniques have been invented and are continuing to be invented to deal with specific aspects of this complexity, for example, science-technology-society issues analysis strategies. At a metalevel, there seems to be some promise of coherence in viewing the matter of curriculum content from an epistemological perspective. There is evidence (Jenkins, 1987) that this can be achieved within the constraints of a conventional teaching setting, particularly if classroom teachers are actively involved in writing the curriculum materials and in developing the teaching strategies to do this.

Schools are expected to fulfill a number of often competing social functions. In the science area, Fensham (1988) identifies two clusters of such functions as:

a. political, economic and subject maintenance;
b. cultural, social and individual.

The potential for curriculum competition, change, and evolution when the varieties of normative perspectives it is possible to take on a domain of human knowledge are juxtaposed with the competing social functions that schooling is to serve suggests that the field of curriculum and instruction will likely be an active one for some time to come in the realm of science education. As a start, we are beginning to identify some of the dimensions of the problem as it seems to exist today.

References

Aikenhead, G., & Fleming, R. (1975). *Science, a way of knowing*. Saskatoon, SK: Department of Curriculum Studies, University of Saskatchewan.

Alberta Education. (1990). *STS Science education: Unifying the goals of science education*. Edmonton, AB: Curriculum Support Branch.

Apple, M. (1979). *Ideology and the curriculum*. London: Routledge and Kegan Paul.

Atlantic Science Curriculum Project. (1990). *Science plus*. Toronto, ON: Harcourt Brace Jovanovich.

Baird, J., & Mitchell, I. (Eds.). (1986). *Improving the quality of teaching and learning: An Australian case study—the Peel project*. Melbourne, Victoria: Monash University Printery.

Bhaskar, R. (1975). *A realist theory of science*. Leeds, UK: Leeds Books.

Bradley, J. (1975). Where does theory begin? *Education in Chemistry, 12*, 8-11.

Bruner, J., Goodnow, J., & Austin, G.F. (1956). *A study of thinking*. New York: Wiley.

Connelly, F.M. (1972). The functions of curriculum development. *Interchange, 3*(2-3), 161-177.

Connelly, F.M., Crocker, R., & Kass, H. (1989). *Science education in Canada, vol. 2. Achievement and its correlates*. Toronto, ON: OISE.

Desautels, J. (1982). What sort of scientific education for what sort of society? *Quebec Science Education: Which Directions?* Ottawa: Science Council of Canada.

Factor, L., & Kooser, R. (1981). *Value presuppositions in science textbooks: A critical bibliography*. Washington, DC: National Science Foundation.

Fensham, P. (1975). Concept formation. In D.J. Daniels (Ed.), *New movements in the study and teaching of chemistry*. London: Temple Smith.

Fensham, P. (1983). A research base for new objectives of science teaching. *Science Education, 67*(1), 3-12.

Fensham, P. (1988). Familiar but different: Some dilemmas and new directions in science education. In P. Fensham (Ed.), *Development and dilemmas in science education*. London: Falmer.

Fensham, P., & Kass, H. (1988). Inconsistent or discrepant events in science instruction. *Studies in Science Education, 15*, 1-16.

Hammond, G., & Nyholm, R. (1971). The structure of chemistry. *Journal of Chemical Education, 43*, 6-13.

Hawking, S. (1988). *A brief history of time*. New York: Bantam.

Jenkins, F. (1987). *A taxonomy of curricular discourse*. Unpublished doctoral dissertation, University of Alberta.

Jenkins, F., Klimiuk, G., Tompkins, D., Hurt, D., Lantz, O., Dzwiniel, M., Kuzul, E., Jackson, D., Mowat, T., & Baziuk, M. (1981). *ALCHEM* (17 vols.) Edmonton: LeBel.

Jenkins, F., Dzwiniel, M., Klimiuk, G., Lantz, O., Tompkins, D., & van Kessel, H. (1989). *Science, technology society, communication, CHEMISTRY*. Edmonton: Author Group Karitann Publication.

Kahle, J.B. (1988). Gender and science education II. In P. Fensham (Ed.), *Development and dilemmas in science education*. London: Falmer.

Kuhn, T. (1970). *The structure of scientific revolutions*. Chicago. University of Chicago Press.

Lemke, J. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex.

Martin, B., Kass, H., & Brouwer, W. (1990). Authentic science: A diversity of meanings. *Science Education, 74*(5), 541-554.

Osborne, R., & Freyberg, P. (1985). *Learning in science. The implications of children's science*. Auckland, NZ: Heinemann.

Osborne, R., & Wittrock, M. (1985). The generative learning model and its implications for science education. *Studies in Science Education, 12*, 59-87.

Pines, L., & West, L. (Eds.). (1985). *Cognitive structure and conceptual change*. New York: Academic Press.

Popper, K. (1959). *The logic of scientific discovery*. London: Hutchinson.

Ravetz, J.R. (1971). *Scientific knowledge and its social problems*. New York: Oxford University Press.

Roberts, D. (1982). Developing the concept of "curriculum emphasis" in science education. *Science Education, 66*(2), 243-60.

Science Council of Canada. (1984). *Science education in Canadian schools: Background study 52*. Hull, PQ: Minister of Supply and Services, Canadian Government Publishing Centre.

Sjoberg, S., & Imsen, G. (1988). Gender and science education I. In P. Fensham (Ed.), *Development and dilemmas in science education*. London: Falmer.

Stake, R.E., & Easley, J. (Eds.). (1978). *Case studies in science education*. Urbana, IL: Center for Instructional Research and Curriculum Evaluation, University of Illinois.

Walker, D. (1971). A naturalistic model for curriculum development. *School Review, 80*, 51-65.

Weiss, I.R. (1978). *Report of the 1977 National Survey of science, mathematics, and social studies education*. Washington, DC: U.S. Government Printing Office, #038-000-0364-0.

Weisskopf, V. (1976). Is physics human? *Physics Today, 29*, 23-29.

White, R.T., & Tischer, R.P. (1986). Research on natural sciences. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). AERA. New York: Macmillan.

The Alberta Journal of Educational Research *Vol. XXXVII, No. 4, December 1991, 385-397*

ROBERT F. MULCAHY
University of Alberta

Developing Autonomous Learners

Recent developments in cognitive psychology and education indicate that students who demonstrate autonomous or self-regulated learning are better equipped to promote their own academic achievement. In this article the concept of autonomous learning is first briefly described. Then an instructional approach to developing autonomous learners is presented including its underlying assumptions, instructional model, and teacher training procedures, as well as some research findings and anticipated future development.

Since about the beginning of the 1960s there has been an increasing interface of cognitive psychology and education. Evidence of this can be found in the emphasis now being placed on the development of learning/thinking strategies in educational assessment and instruction. The recent upsurge in interest in autonomous learning or self-regulation (McCombs, 1989; Paris & Byrnes, 1989; Whitman, 1990; Zimmerman & Schunk, 1989) is a result of this movement.

A number of papers are available in the literature addressing the issue of self-regulation and autonomous learning (see, e.g., Zimmerman & Schunk, 1989). This article briefly discusses the concept of self-regulation or autonomous learning and then outlines an approach to instruction we have been working on at the University of Alberta for the past 10 years. The approach, *Strategies Program for Effective Learning/Thinking* (SPELT), has as its goal the development of an autonomous learner.

The Concept of Autonomous Learning

The key feature of what makes an autonomous learner appears to be the independent self-control of one's cognitive as well as affective resources and activities. This position is supported by a number of researchers and theorists. For example, Paris and Byrnes (1989) conceive the self-regulated learner as one who seeks challenges, overcomes obstacles, sets realistic goals, and utilizes a battery of resources in approaching tasks with confidence and purpose. They see the hallmark of a self-regulated learner as combining positive expectations, motivation, and a wide range of strategies for solving problems. The view of the child as "theorist" is emphasized. According to this view, the child's theories of self-competence, effort, and academic tasks interact to develop or inhibit self-regulated learning.

Similarly, Palinscar, David, Winn, and Stevens (1991) depict the self-regulated or autonomous learner as one who uses with flexibility three main kinds of knowledge: knowledge of strategies for performing learning tasks

Robert Mulcahy is a professor in the Department of Educational Psychology. His research interests focus on the area of metacognition, cognitive strategy assessment, and instruction.

efficiently; metacognitive knowledge of one's own learner characteristics and task demands; and real world knowledge. Moreover, they implicate motivation as they suggest that the self-regulated learner not only utilizes the above three knowledge bases, but as well demonstrates the motivation to employ the knowledge effectively.

McCombs (1989) also views the self-regulated learner as a function of the development of cognitive and metacognitive abilities for processing, planning, and regulating learning activities, as well as for controlling and regulating affect and motivation. McCombs, however, explicitly places greater emphasis on the importance of a positive affect and motivation in the emergence of the autonomous learner. Some other researchers (e.g., Dweck, 1986; Skinner, Chapman, & Baltes, 1988) also contend that students' beliefs, expectations, and attributions with respect to academic tasks have a significant impact on their learning and indeed on whether or not they are able to become autonomous learners .

Self-regulated or autonomous learning, then, appears to mean that individuals are able to independently manage their learning and affect including motivation, so that learning is efficient, effective, and satisfying (Paris & Byrnes, 1989).

In contrast to this "ideal" model of the autonomous learner, there is a relatively large data base in the literature characterizing ineffective learners as those who often fail to demonstrate the flexible use of various knowledge bases as well as failing to display positive belief systems with respect to their personal competence and control of the learning process (cf. Paris & Oka, 1986). As a result, numerous instructional programs aiming to improve students' learning and thinking have been developed (cf. Nickerson, Perkins, & Smith, 1985).

The SPELT Cognitive Education Approach

For the past 10 years we have been involved in the development and evaluation of an instructional approach to teaching metacognitive and cognitive strategies along with the control and regulation of affect and motivation. This instructional approach is called *Strategies Program for Effective Learning/Thinking* (SPELT). In the following sections, the underlying assumptions of the approach, its instructional methods, teacher training procedures, and research findings and future development are described.

Underlying Assumptions

The underlying assumptions of our approach are that learning is an active, constructive process that involves the individual's use and management of his or her own cognitive processes. As the overall goal of the approach is to develop a self-regulated or autonomous learner, the instructional process utilized is of major importance.

The dominant model in the majority of cognitive and metacognitive instruction approaches, however, has been the instructional system assigned or teacher imposed. Typically, the strategies are designed and tested by experts and then taught to students as recipes for dealing with a variety of problems. Within this framework there is virtually no opportunity for students to participate in determining which cognitive or metacognitive strategies are

appropriate for different purposes and varying conditions. This approach also reinforces a passive learning approach as well as reinforcing an external locus of control, both of which run counter to the development of an autonomous or self-regulating learner.

In contrast, the SPELT approach we have been developing utilizes a model of instruction in which the student is ultimately given the responsibility of determining if a strategy is needed, what strategies are appropriate, and how to generate, implement, and evaluate them in a way that maximizes problem solution.

Cognitive and metacognitive strategies are viewed in our approach as internally organized skills or control processes by which individuals regulate their behavior (both affective and cognitive). An individual's knowledge base regarding cognitive and metacognitive strategies and procedures for utilizing these along with the knowledge of what, when, where and why to utilize a particular strategy or combination of strategies is seen as the vehicle by which the individual may become a self-regulated learner.

Within the SPELT approach we have imbedded a general teaching orientation in which the teacher's goal in all planning and instruction is to actively involve the student in his or her learning. The teacher acting as a mediator in the learning process attempts to have students become more aware of their own cognitive and metacognitive processes as well as aware of their motives and affect as it affects their learning. In acting as a mediator the teacher leads students to discover and deduce rather than teaching facts to them, as well as constantly challenging students to be more critical, systematic, evaluative, and strategic in their behavior and attitude toward learning.

Instructional Methods

How can the teacher be successful in acting as a mediator? We incorporated three major types of instructional methods into the program to nurture the students' learning and motivation that would lead to self-regulated learning (Mulcahy, Marfo, & Peat, 1984). The first type is direct instruction and explanation of learning strategies that may help children modify incomplete or erroneous theories they may hold with respect to self-regulated learning. The second type of instruction emphasizes peer tutoring and dialogues about learning. Instruction here also may involve Socratic dialogues, apprenticeship models, and dialogical learning. The third category of instruction involves cooperative learning involving group discussion, argument, and co-construction of appropriate cognitive and metacognitive strategies. These three types of instruction have recently been suggested to be effective innovative methods to assist teachers in developing self-regulated learners (Paris & Byrnes, 1989).

With these three types of instruction integrated into our program, the instructional approach involves a progression from the lowest level of strategy acquisition (teacher imposed) to the highest level (student self-generated). Accordingly, we have developed three phases of instruction in which the teacher taking the role of a mediator assists the students in becoming increasingly more in control of the learning process leading them ultimately to autonomous learning.

The first phase of instruction involves the direct teaching of cognitive and metacognitive strategies that are program recommended, teacher identified, or teacher generated. This phase of the instruction approach has as its goals the development of a repertoire of metacognitive and cognitive strategies, along with making students more aware of the benefits of strategic behavior. As well an emphasis in the instruction here is on developing positive perceptions of self-competence along with attributional beliefs that emphasize personal control and competence. During this phase of instruction the source of control of the strategies resides with the instructor.

Some researchers have criticized cognitive strategy instruction approaches as involving the mere memorization and rote learning of strategy steps and application (cf. Kronick, 1988). If instruction were to remain at this first level, it is possible that it would encourage many of the negative characteristics of learning such as passivity and external locus of control. However, it should be noted that "recent research ... indicates that the process of knowledge transformation and construction occur even when children engage in drill-and-practice activities" (Harris & Pressley, 1991, p. 393). More importantly, our approach includes two more advanced phases of instruction.

The second phase of instruction in the SPELT approach is what might be considered the transfer and generalization phase. Instruction here emphasizes systematic transfer and generalization of the strategies built in the first phase of instruction across various tasks and content areas. The major mode of instruction moves to a more dynamic interactive and reflective one in which the teacher relies heavily on Socratic dialogues to stimulate students to reflect and reconsider their views, as well as encouraging them to develop explicit understanding of tasks, strategies, and their own motives, feelings, and beliefs regarding these. The teacher is encouraged to utilize a number of other methods in conjunction with this dialogue including paired and group problem solving, thinking aloud, cooperative group, brainstorming on the pros and cons of various cognitive and metacognitive strategies for different tasks, situations, and purposes. The students are led through teacher and group dialogue to become more aware of their own task approach and strategy use as well as others, and use this to critically analyze and then extend or modify their own personal strategies or the strategies presented to them.

By the time students are functioning in the third phase (student self-generation of strategies) of instruction, they have developed a knowledge base regarding various cognitive and metacognitive strategies along with knowledge of their procedures for application, including an awareness of where, when, and why they should be used. This then serves as the knowledge base for student self-generated strategies. During this phase of instruction content is presented to the students with minimal amounts of teacher guidance as to how to complete the task. The teacher, utilizing the Socratic dialogue, centers discussion around an analysis of how the task was completed or is being approached by different students, analyzing of task requirements, alternative approaches, and ways of determining the relative effectiveness of

possible strategies. Here again, paired and group problem solving thinking aloud are recommended.

A variety of cognitive and metacognitive strategies are integrated into the approach, including strategies that are effective for retaining information, comprehending information, communicating information (written and spoken), controlling affect and motivation, general and specific problem solving, and numerous others (see Peat, Mulcahy, & Darko-Yeboah, 1989, for a listing of strategies taught).

Inservice Training Procedures

The successful implementation of SPELT depends largely on the teacher. Therefore, inservice training is an integral part of the approach. The inservice training of teachers was originally conducted over a five-full-day session broken into two blocks. Part One was a three-day session and was followed by Part Two about three weeks later.

During Part One training, teachers are introduced to the program's theoretical framework in cognitive psychology, to the three-phase instructional model, as well as the practical applications of the model. Throughout Part One training teachers are continually exposed to cognitive theory of learning. Teachers are instructed to focus on the development of positive self-competence and attributional belief systems of the students in conjunction with instruction in cognitive and metacognitive strategies.

Part Two Training has as its major emphasis the systematic transfer of students' strategic repertoire established earlier to other situations, settings, and applications (i.e., generalization). As the applications of the various strategies change, teachers are instructed to discuss with their students these strategy adaptations, modifications, and/or extensions, emphasizing how these changes effect their use. Students thus see that once a strategy is learned it is *not* carved in stone, but can be modified. This personalizes the strategies and acts as a stepping stone in the students' ability to self-generate effective learning and thinking strategies: the goal of the final phase of instruction.

During Part Two training, teachers are instructed in the principles and use of *Socratic dialogue*, which involves an interactive relationship between teacher and students with the students being led through questioning to discover for themselves. Socratic dialogue is operationalized based on the work of Collins (1977), specifying guidelines such as: (a) starting with what is know, (b) asking for multiple reasons, (c) forming general rules from specific cases, (d) picking counter examples when insufficient reason is given, (e) using extreme case examples, (f) probing for the difference between cases, and (g) using prediction questions. Key teacher behaviors that facilitate Socratic dialogue are also recommended (e.g., 5-10-second wait time). It is important to note that Socratic dialogue is the chosen methodology for discussion centering around the strategy use of students, but it becomes readily apparent to teachers how this type of dialogue can enhance the thinking skills of students during any classroom discussion, regardless of content.

As the teachers recognize some of the specific strategies or methodologies presented, they begin to see that some aspects of their past classroom teaching has emphasized the teaching of learning/thinking skills, but that instruction in this area has usually been unconscious and/or incidental, lacking a precise goal or direction. This recognition aids in their conceptualizing "cognitive education," in spite of their antitheory bias, as nonthreatening and achievable in their classrooms. Also, during the workshop the teachers themselves actively practice the strategies and engage in paired problem solving, cooperative learning, and Socratic dialogues by being placed in the role of students as the presenter's instructional approach moves along the SPELT continuum.

This model allows the teachers to experience first hand the *process* that parallels what their students will undergo as the approach is implemented in their classrooms. The process allows the teachers to freely express their doubts and concerns throughout the inservice, a factor shown to be important in facilitating receptivity to change. Placing the teachers in the student position also has the effect of concretely demonstrating to them that they too should continue to develop their thinking skills and, in that sense, places them in the dual role of learner and teacher. It also serves to emphasize their responsibility as role models of systematic and strategic problem solvers.

We have over the past seven years experimented with a variety of inservice training formats for the instruction of teachers, counselors, psychologists, and administrators. The optimum appears to be three days of intensive instruction followed in about one month by one full day and then two months later another full day.

A Longitudinal Study

The SPELT cognitive education approach has been extensively evaluated in a three-year longitudinal study jointly supported by Alberta Education, Government of Alberta; Faculty of Education, University of Alberta; and school jurisdictions in Alberta. The detailed results of this study have been reported elsewhere (Mulcahy et al., 1989; Mulcahy, et al., 1991). The study involved two years of experimental instruction followed by a one-year maintenance period where all instruction was withdrawn. Nine hundred students (learning difficulties, average, and gifted) initially in regular grade 4 and grade 7 experimental and control classrooms were followed over the course of three years. The overall results of the study were very encouraging both from the perspective of student change as well as from the perspective of the teachers, parents, and administrators.

The three-year longitudinal evaluation was implemented in two phases (i.e., phase I: 1984-1987; and phase II: 1985-1988). We utilized a repeated measures factorial design involving three types of instructional programs (Instrumental Enrichment, SPELT, and Control); three categories of students (gifted, average, and learning-disabled); two initial grade levels (grade 4 and grade 7), and four repeated measures: pretest in the fall of the initial year, and three posttests in succeeding May/June periods, corresponding to the end of grades 4, 5, and 6 at the elementary level and 7, 8, and 9 at the junior high level.

Teachers volunteered for any one of the three instructional conditions. The research team then randomly assigned them to an instructional condition with the constraint that a given school could only be involved with one instructional condition.

Teachers assigned to the control condition (traditional instruction) were told to teach as usual, whereas teachers assigned to the two cognitive education procedures received inservice training from project staff prior to giving strategy instruction. Students involved in the two cognitive education programs received a minimum of 120 minutes of strategy instruction per week over two school years. Strategy instruction was followed by one year of maintenance, during which all strategy instruction was withheld.

Four categories of instruments were utilized in this study (full details in Mulcahy et al., 1989):

1. *Cognitive ability* was assessed by the Canadian Cognitive Abilities Test (CCAT), which includes verbal, quantitative, and nonverbal reasoning abilities.
2. *Academic achievement* was assessed by the Canadian Achievement Test (CAT), which consists of two separate batteries measuring skills in reading (vocabulary, comprehension) and mathematics (computation, concept/application).
3. *Affective perceptions* involved measures of perceived competence in four areas (cognitive, social, physical, and general); attitudes toward the self in social, academic, family, and personal areas of experience; and students' beliefs regarding responsibility for outcomes in academic achievement situations.
4. *Cognitive strategies* were assessed in the areas of:
 a. *Reading awareness*, which included students' awareness of the evaluation, planning, and regulation of skills involved in reading.
 b. *Reading strategies*. A cloze task was developed to evaluate students' reading-comprehension abilities and strategies; words were systematically omitted, and students required to fill in the blanks.
 c. *Comprehension monitoring*. An error detection task, designed to assess students' comprehension monitoring skills, required students to detect anomalous information in the passage. Awareness of faulty comprehension is taken as a measure of students' comprehension monitoring.
 e. *Perceived problem solving ability*. Heppner and Petersen's (1982) Problem-Solving Inventory (PSI) was adopted to examine the underlying dimensions of students' perception of their real-life, personal problem-solving process.
 f. *Problem solving strategies*. A problem solving task was developed to assess students' problem solving strategies in mathematics. This task requires students to think aloud as they try to solve the given math problems.

Most of the above measures were implemented in both the pretest and posttest phases, with the exception that measures of cognitive strategies were not available at the pretest point and were administered at the posttest points only.

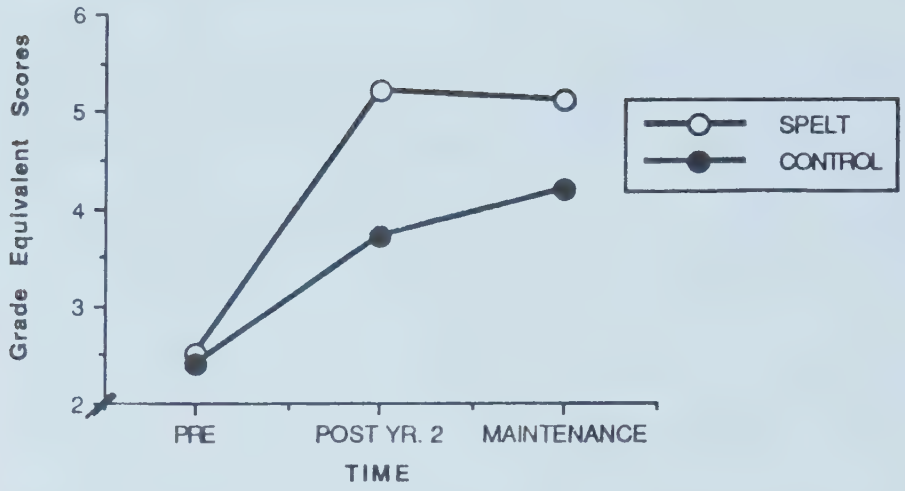


Figure 1. Reading comprehension: Grade 4 learning disabled.

Teacher, administrator, and parent perceptions of program implementation and effects were also evaluated each year.

Research Findings

Because of the large amount of data, only the results with respect to the SPELT approach are reported here. To provide some idea of some particular changes observed, a brief summary of the results for reading comprehension performance, metacognitive awareness in reading, comprehension monitoring, along with teachers' perceptions are described in the following section. More complete results are provided in Mulcahy et al. (1989), including statistical details.

Grade 4. The grade 4 learning disabled SPELT students displayed greater gains than their control counterparts on the standardized measure of reading comprehension after two years of instruction as this difference tended to be maintained at maintenance (see Figure 1).

At the end of the three years, the SPELT students were performing at about one grade level higher than control subjects (scaled score analysis demonstrated the same effects). These changes were coupled with changes in strategy areas of reading, particularly with respect to metacognitive reading awareness (see Figure 2).



Figure 2. Metacognitive reading awareness: Grade 4 learning disabled.

The average achieving students demonstrated few experimental effects on the standardized measures of reading comprehension; however, some significant group differences were obtained with respect to metacognitive reading awareness, the SPELT average students generally outperforming their control counterparts at all test points over the three years. This was the case as well with respect to comprehension monitoring (see Figure 3).

Gifted students in the SPELT program did not differ from their control counterparts on standardized measures of reading. There was indication, however, that the SPELT approach did have a positive impact on metacognitive awareness in reading as differences were noted at all three test points in favor of SPELT subjects. This difference was also observed for the comprehensive monitoring performance of the two groups.

Grade 7. A trend was observed for differences in gains in reading comprehension in favor of the SPELT students as measured by the standardized test of reading (see Figure 4). The trend was also evident with respect to the standardized measure of vocabulary (see Figure 5).

No significant differences were obtained with respect to metacognitive awareness in reading for the grade 7 SPELT group. However, the overall group mean followed the same pattern as that observed at grade 4 (SPELT

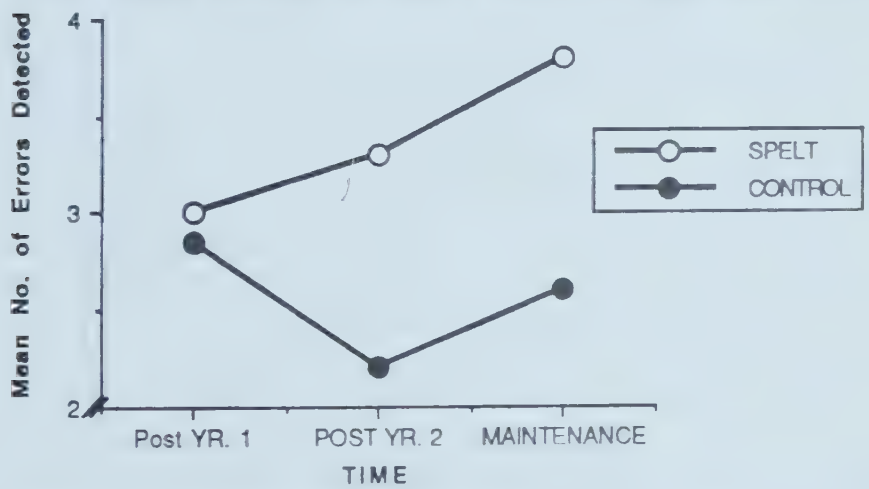


Figure 3.Comprehension monitoring: Grade 4 average students.

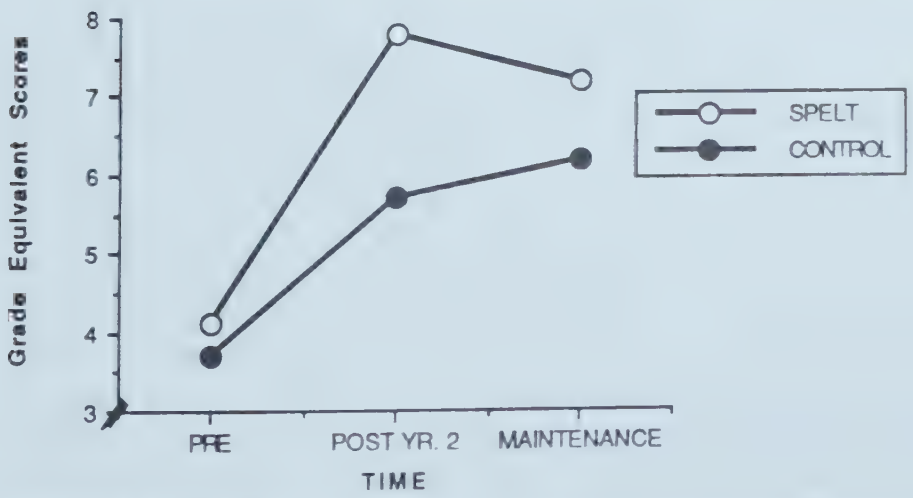


Figure 4. Reading comprehension: Grade 7 learning disabled.

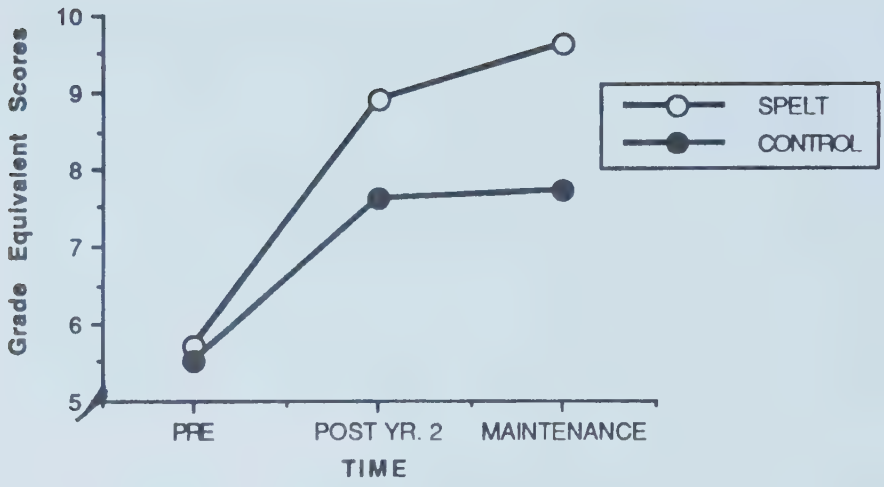


Figure 5. Reading vocabulary: Grade 7 learning disabled.

mean raw score= 33.4, Control mean raw score= 30.9). Overall differences were also obtained with respect to comprehension monitoring in favor of the SPELT students (SPELT mean raw score = 3.6, Control mean raw score = 2.7).

Average achieving grade 7 students in the SPELT approach did not differ significantly from control students on standardized measures of reading or on measures of comprehension monitoring and metacognitive awareness.

Gifted grade 7 students in the SPELT approach did not differ from control students on standardized measures of reading. Nevertheless, differences in favor of the SPELT students were observed on the measure of metacognitive awareness in reading (see Figure 6). The SPELT students also outperformed control students on the measure of comprehension monitoring, particularly at the end of the three years (see Figure 7).

In summary, the SPELT results for measures of reading comprehension and vocabulary, as well as metacognitive awareness in reading and comprehension monitoring, are encouraging. The lack of effects observed for grade 7 average achieving and to some extent grade 4 average students is somewhat puzzling. However, as we have argued elsewhere (Mulcahy et al.,

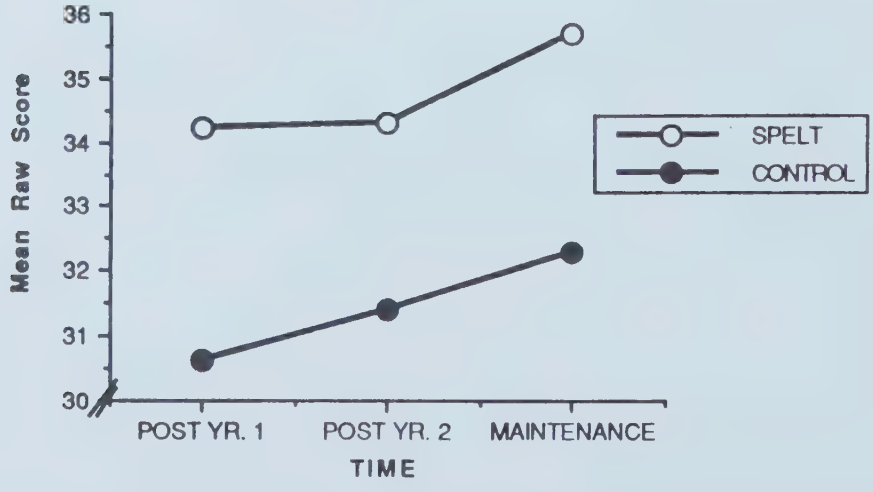


Figure 6. Metacognitive reading awareness: Grade 7 gifted.

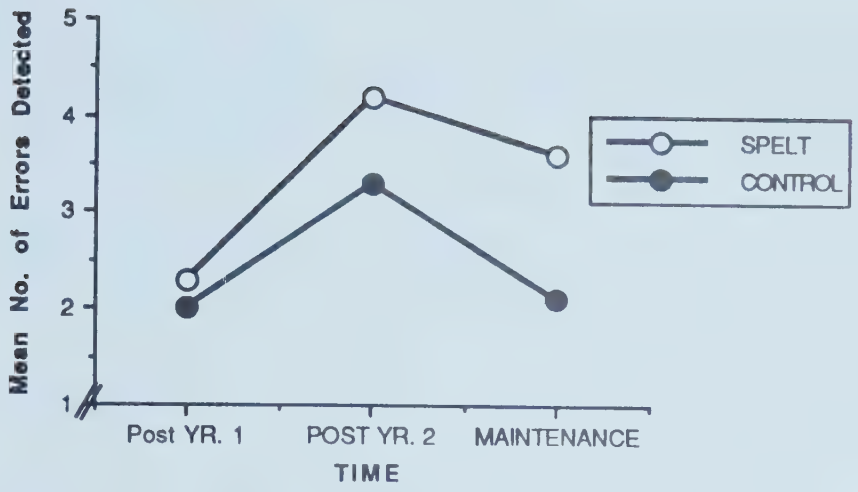


Figure 7. Comprehension monitoring: Grade 7 gifted.

1991), it might well be because average achieving students have a reasonably effective approach to tasks already and thus a greater effort is needed to change or refine their strategic approach, so they do not benefit from the first phase of instruction. Learning disabled students, on the other hand, may lack a systematic approach to tasks and thus benefit more quickly. Gifted students may perceive the usefulness of a more refined approach and thus use and extend the strategies more quickly, thus benefiting more from the final phase of instruction.

The above results by themselves have relatively clear implications for mainstreaming learning disabled and gifted students. The change noted here in the SPELT approach for both gifted and learning disabled students suggests that this instructional approach in regular classrooms may prevent some students from developing severe learning problems, therefore enabling them to function effectively in regular classrooms. The gifted students, as identified in this study, also appeared to benefit from this instruction in regular class settings.

This particular evaluation was quite conservative, as in the majority of situations only one or two teachers in a school were teaching using the approach at one grade level and in junior high usually in one subject level. If *all* teachers at each grade level were focusing on instruction in metacognitive and cognitive strategies, more comprehensive effects might emerge as students would be reinforced across grades and subjects.

In addition to the above experimental findings, it should be pointed out that the inservice training sessions have also been consistently evaluated over the course of the years with encouraging results. We have now inservice trained approximately 2,000 teachers, counselors, psychologists, speech language pathologists, and school administrators in Canada, Australia, and New Zealand. Teachers not only recommend this type of instructional training to their colleagues, but also a large number have suggested that this approach to classroom instruction be an integral part of the preservice training for all teachers. Many indicate that the approach has changed their teaching style and the way they view their students and the learning process.

Current Development

We are currently focusing on within-classroom, within-subject research in an attempt to delineate more clearly the problems and issues in implementation, as well as to determine the essential components of instruction that are critical to developing self-regulated learners in the classroom.

Teachers' perceptions is another area of further research we are planning to pursue in order to determine the effect this approach has on teachers' view of development, learning, and the teaching process as well as the nature of the classroom interaction patterns.

Finally, the implementation and evaluation of the SPELT approach is now being extended to cross-cultural research sites in Australia, New Zealand, and India. We are also currently exploring the further development of the approach with aboriginal students in Australia and Canada. The continued development and validation of assessment tools and approaches for metacognitive and cognitive strategy assessment within and across domains, including the social and affective areas, is an integral part of this further research. The longstanding goal of our research is to add to our knowledge regarding the nurturing of autonomous learning in a variety of contexts.

Acknowledgments

The author would like to acknowledge the assistance and support of the co-workers in the SPELT approach: Kofi Marfo, David Peat, and Joe Andrews.

References

Collins, A. (1977). Processes in acquiring knowledge. In R.C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 339-363). Hillsdale, NJ: LEA.

Dweck, C.S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040-1048.

Harris, K.R., & Pressley, M. (1991). The nature of cognitive strategy instruction: Interactive strategy construction. *Exceptional Children, 57*(5), 393-403.

Heppner, P.P., & Peterson, C.P. (1982). The development and implications of a personal problem-solving inventory. *Journal of Counselling Psychology, 29*, 66-75.

Kronick, D. (1988). *New approaches to learning disabilities: Cognitive, metacognitive, and holistic*. Philadelphia, PA: Grune and Stratton.

McCombs, B.L. (1989). Self-regulated learning and academic achievement: A phenomenological view. In B.J. Zimmerman & D.H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research and practice* (pp. 51-82). New York: Springer-Verlag.

Mulcahy, R., Andrews, J., & Peat, D. (1989). Cognitive education: A longitudinal evaluation. In C.K. Leong & B.S. Randhawa (Eds.), *Literacy and Cognition: Theory, Research and Instructional Implications* (pp. 287-314). New York: Plenum.

Mulcahy, R., & Marfo, K. (1987). Assessment of cognitive ability and instructional programming with Native Canadian children: A cognitive processes perspective. In L. Stewin & S. McCann (Eds.), *Contemporary issues in education: A Canadian mosaic* (pp. 157-177). Toronto, ON: Copp Clarke-Pittman.

Mulcahy, R., Marfo, K., & Peat, D. (1984). *SPELT: A strategies program for effective learning and thinking. Research edition* (available from the Cognitive Education Project, Department of Educational Psychology, University of Alberta).

Mulcahy, R., Peat, D., Andrews, J., Darko-Yeboah, J., Marfo, K., & Cho, S. (1989). *Cognitive Education Project: Final Report*. Edmonton, AB: Alberta Education.

Mulcahy, R., Peat, D., Andrews, J., Darko-Yeboah, J., Marfo, K., & Cho, S. (1991). Cognitive strategy based instruction. In J. Biggs (Ed.), *Teaching for learning: The view from cognitive psychology* (pp. 193-214). Sydney, Australia: ACER.

Nickerson, R., Perkins, D., & Smith, E. (1985). *Teaching thinking*. New York: Academic Press.

Palinscar, A.S., David, Y.M., Winn, J.A., & Stevens, D.D. (1991). Examining the context of strategy instruction. *Remedial and Special Education, 12*(3) 43-53.

Paris, S.G., & Byrnes, J.P. (1989). The constructivist approach to self-regulation and learning in the classroom. In B.J. Zimmerman & D.H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research and practice*. New York: Springer-Verlag.

Paris, S.G., & Oka, E.R. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review, 6*, 25-56.

Peat, D., Mulcahy, R.F., & Darko-Yeboah, J. (1989). SPELT (Strategies Program for Effective Learning/Thinking): A description and analysis of instructional procedures. *Instructional Science, 18*, 95-118.

Skinner, E.A., Chapman, M., & Baltes, P.B. (1988). Control, means-ends, and agency beliefs: A new conceptualization and its measurement during childhood. *Journal of Personality and Social Psychology, 54*, 117-133.

Whitman, R.L. (1990). Self-regulation and mental retardation. *American Journal on Mental Retardation, 94*(4), 347-362.

Zimmerman, B.J., & Schunk, D.H. (Eds.). (1989). *Self regulated learning and academic achievement: Theory, research, and practice*. New York: Springer-Verlag.

Index, Volume XXXVII, 1991

Beck, D. *Redefining Research Relationships: Two Heads are Better than One*. No. 2, June, 133-140.

Beebe, M.J. *Gender Differences, Student Well-Being and High School Achievement*. No. 3, September, 209-224.

Black, K. *Redefining Research Relationships: Two Heads are Better than One*. No. 2, June, 133-140.

Bulcock, J.W. *Gender Differences, Student Well-Being and High School Achievement*. No. 3, September, 209-224.

Campbell-Evans, G.H. *Nature and Influence of Values in Principal Decision Making*. No. 2, June, 167-178.

Chamberlin, C. *Reflections on a Collaborative School-Based Teacher Education Project*. No. 2, June, 141-156.

Chow, H. *Issues in Implementing Indicator Systems*. No. 1, March, 65-86.

Covert, J. *Some Perceived Professional Needs of Beginning Teachers in Newfoundland*. No. 1, March, 3-17.

Erickson, G. *Prior Experience and Gender Differences in Science Achievement*. No. 3, September, 225-239.

Echols, F. Jr. *Students' Knowledge of the Charter of Rights and Freedoms and Their Agreement With Its Provisions*. No. 1, March, 31-48.

Farkas, S. *Prior Experience and Gender Differences in Science Achievement*. No. 3, September, 225-239.

Haggerty, S. *Gender and School Science: Achievement and Participation in Canada*. No. 3, September, 195-208.

Hague, W.J. *Kohlberg's Legacy—More Than Ideas. An Essay Review*. No. 3, September, 277-294.

Hattie, J. *The Burt Controversy: An Essay Review of Hearnshaw's and Joynson's Biographies of Sir Cyril Burt*. No. 3, September, 259-275.

Hodysh, H.W. *A Note on History, Educational Policy, and the Uses of the Past*. No. 4, December, 323-332.

Jacknicke, K.G. *Some Perceptions of an Internship Program*. No. 2, June, 99-118.

Kass, H. *The Transformation of Curricular Knowledge*. No. 4, December, 367-383.

Kennedy, W. *Some Perceived Professional Needs of Beginning Teachers in Newfoundland*. No. 1, March, 3-17.

La Bar, C. *Students' Knowledge of the Charter of Rights and Freedoms and Their Agreement With Its Provisions*. No. 1, March, 31-48.

MacDonald, C.J. *Elementary School Teachers Explain Why They Do Not Use Creative Dance in Their Classrooms*. No. 2, June, 157-166.

Malicky, G.V. *Myths and Assumptions of Literacy Education*. No. 4, December, 333-347.

McEwen, N. *Issues in Implementing Indicator Systems*. No. 1, March, 65-86.

Miklos, E. *Trends in Doctoral Research in Educational Administration*. No. 4, December, 297-321.

Mulcahy, R.F. *Developing Autonomous Learners*. No. 4, December, 385-397.
Oberle, K. *Paradigm Wars: Who's Fighting; Who's Winning?* No. 1, March 1991, 87-97.
Olson, M.R. *Supportive Growth Experiences of Beginning Teachers' Practices With Regard to Oral Language*. No. 1, March, 19-30.
Randhawa, B.S. *Gender Differences in Academic Achievement: A Closer Look at Mathematics*. No. 3, September, 241-257.
Rogers, W.T. *Educational Assessment in Canada: Evolution or Extinction?* No. 2, June, 179-192.
Rogers, W.T. *Introduction to Special Section on Gender Differences in Achievement*. No. 3, September, 193-194.
Samiroden, W.D. *Some Perceptions of an Internship Program*. No. 2, June, 99-118.
Sawada, D. *Deconstructing Reflection*. No. 4, December, 349-366.
Shapiro, B.L. *A Collaborative Approach to Help Novice Science Teachers Reflect on Changes in Their Construction of the Role of Science Teacher*. No. 2, June, 119-132.
Staab, C. *Teachers' Practices With Regard to Oral Language*. No. 1, March 1991, 31-48.
Ungerleider, C.S. *Students' Knowledge of the Charter of Rights and Freedoms and Their Agreement With Its Provisions*. No. 1, March, 31-48.
Vallance, J. *Reflections on a Collaborative School-Based Teacher Education Project*. No. 2, June, 141-156.
Whitt, M.E. *Gender Differences, Student Well-Being and High School Achievement*. No. 3, September, 209-224.
Williams, L. *Some Perceived Professional Needs of Beginning Teachers in Newfoundland*. No. 1, March, 3-17.
Zingle, H. *You've Come a Long Way!* Guest Editorial. No. 4, December, 295-296.

INTERCHANGE

A QUARTERLY REVIEW OF EDUCATION

- a leading international journal on the critical issues facing educators in Canada and abroad
- a recipient of the prestigious CAFE award for the most original work in educational foundations
- four issues per year feature scholarly articles, critical notices, book reviews, and debates
- recent special issues include Creativity, Thought, and Mathematical Proof; History, Science and Science Teaching; and Understanding Literacy (guest edited by David Olson)
- Edited by Ian Winchester

Regular One-Year Rates for 1990: Institutions $50.00; Individuals $35.00. For 1991: Institutions $55.00; Individuals $35.00. To order send cheque, purchase order, or complete MasterCard or VISA information to OISE Publishing, 252 Bloor St. W., Toronto, Ontario, Canada M5S 1V5.



Published by The Ontario Institute for Studies in Education, Toronto, Canada

If undelivered return to:

The Alberta Journal of Educational Research
4-116 Education North
University of Alberta
Edmonton, AB
T6G 2G5

Second Class Mail Registration Number 1438